12-7-2018 1:30 PM

Multiwavelength, Machine Learning, and Parallax Studies of X-ray Binaries in Three Local Group Galaxies

Robin Arnason
The University of Western Ontario

Supervisor
Barmby, Pauline
The University of Western Ontario

Graduate Program in Astronomy
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Recommended Citation

Arnason, Robin, "Multiwavelength, Machine Learning, and Parallax Studies of X-ray Binaries in Three Local Group Galaxies" (2018). *Electronic Thesis and Dissertation Repository*. 5931.
https://ir.lib.uwo.ca/etd/5931

Abstract

X-ray binary stars are rare systems consisting of a black hole or neutron star and a main-sequence companion star. They are useful probes of galaxy properties and interesting laboratories for extreme physical conditions. In this thesis, I investigated the X-ray binary population of three galaxies in the Local Group.

The Sculptor Dwarf Spheroidal Galaxy offers the chance to study a primordial low-mass X-ray binary (LMXB) population in an isolated, low-metallicity environment. Combining X-ray, optical, and infrared observations, I have studied nine previously identified and discovered four additional LMXB candidates in this galaxy. Of these candidates, all but one are either background galaxies or foreground stars, meaning that Sculptor is presently effectively devoid of bright LMXBs. If Sculptor is able to retain primordial LMXBs at a similar rate to globular clusters, it is likely that bright XRBs in globular clusters observed in the present day were dynamically formed.

The Andromeda Galaxy has the largest catalogue of *Chandra*-studied X-ray sources of any nearby galaxy. I have used this population to test a proof-of-concept method for identifying X-ray binary candidates using machine learning algorithms trained on known sources. After testing a variety of commonly used algorithms, I find that the best-performing random forest algorithm can identify X-ray binary candidates with $\sim$ 85% accuracy. I have identified 16 new strong X-ray binary candidates and find that 4 sources classified as X-ray binaries by this method coincide with star clusters identified by the Panchromatic Hubble Andromeda Treasury project.

The Milky Way's X-ray binary population is the easiest to study but the most challenging for which to accurately measure distance. I have crossmatched Galactic X-ray binary catalogs to the second data release of the *Gaia* mission, finding candidate counterparts for 86 Galactic X-ray binaries. Distances to Gaia candidate counterparts are systematically smaller than those measured using Type I X-ray bursts, suggesting that these bursts do not consistently reach the Eddington limit. High-mass X-ray binaries are correlated with the Galaxy's spiral arms and low-mass X-ray binaries are anti-correlated with the Galaxy's spiral arms at a low level of significance.

Keywords: binaries: X-ray; Galaxies: Local Group; Galaxies, individual: Sculptor Dwarf Spheroidal, M31, Milky Way; stars: black holes, neutron; X-rays: bursts – Galaxy: structure; techniques: parallaxes, methods: statistical, observational

Co-Authorship Statement

All research is the result of collaboration; this thesis is no exception. Each of the three projects presented in this thesis was led by me, and I performed the analysis. I have benefited greatly from guidance, suggestions, and assistance in interpretation from my co-authors: my supervisor Dr. Pauline Barmby, Dr. Arash Bahramian, Dr. Steve Zepf, Dr. Thomas Maccarone, Dr. Neven Vulic, Hadi Papei, and Dr. Mark Gorski.

In **Chapter 2**, Dr. Maccarone and Dr. Zepf were the PIs of the original observations that made the project possible. They also provided assistance in interpretation. Dr. Barmby performed the reduction of the *Spitzer* data - I performed my analysis on the completed *Spitzer* catalogues. Dr. Bahramian took the *Chandra* observations that I reduced and ran them in `ACISEXTRACT`. I performed the analysis and interpretation of the *Chandra* photometry and spectroscopy that are the products of `ACISEXTRACT`. Dr. Barmby and Dr. Bahramian also assisted in interpretation of the results. This paper is submitted to *Monthly Notices of the Royal Astronomical Society*.

In **Chapter 3**, Dr. Vulic created the *Chandra* catalogs used in the study. I joined the individual catalogs for each field, and performed all of the analysis. Dr. Vulic and Dr. Barmby provided guidance both on the catalogue itself, as well as interpretation of the results. This paper will be submitted shortly.

In **Chapter 4**, Dr. Barmby and Mr. Papei performed the work of investigating the method used in the literature to measure distance to each X-ray binary in our matched sample. I cross-matched the catalogues to *Gaia*, though Mr. Papei performed the query to the Bailer-Jones catalogue to obtain the *Gaia* distance to each object. I constructed the simulation and performed the analysis. Dr. Barmby, Dr. Gorski, and Mr. Papei all provided guidance on the interpretation of results. This paper will be submitted shortly.

Epigraph

“The popular stereotype of the researcher is that of a skeptic and a pessimist. Nothing could be further from the truth! Scientists must be optimists at heart, in order to block out the incessant chorus of those who say ‘It cannot be done.’ ”
- *Sid Meier’s Alpha Centauri* (1999)

Dedication

For Mom, Dad, Holly, Ba, and all the birds.

Acknowledgements

This thesis has, like many, been a mix of exciting discovery and keyboard-breaking frustration. No one knows this better than my supervisor, Dr. Pauline Barmby. For the last four years you've been my mentor, my ally, my collaborator, and my friend. You taught me by example that being a good scientist isn't just a matter of doing good research – we also have a duty to make a positive impact on our community, whether on campus or elsewhere. I could not have had a better supervisor.

The direction of my thesis has benefited immensely from guidance from my Advisory Committee members. Thanks to Dr. Jan Cami and Dr. Giovanni Fanchini, who gave me useful research/career advice and struck a balance between pushing me forward and letting me work out the details.

Special thanks also to the collaborators who I have had the privilege of working with over the course of this degree: Dr. Arash Bahramian, Dr. Tom Maccarone, Dr. Steve Zepf, Dr. Neven Vulic, Dr. Mark Gorski, and Hadi Papei. Without your guidance and assistance my thesis would not have gotten off the ground.

Thanks to the wonderful staff and faculty at Western P&A who have been there to help in matters academic and bureaucratic: Clara Buma, Brian Davis, Phin Perquin, Henry Leparskas, Dr. Sarah Gallagher, Dr. Martin Houde, Dr. Els Peeters, Dr. Stan Metchev, Dr. Paul Wiegert, Dr. Silvia Mittler, and Dr. Andy Pon.

Good friends are rarer than X-ray binaries, and I've been blessed to have many here at Western. Keegan, Moh, Dan, Megan, Amanda, Kendra, Isabelle, Hadi, Ghazal, Sina, Sahar, Laura, Aycha, Shannon, Amgad, Jeff, Cameron, and Ellie. Thanks for all the great times and for your support, I couldn't have done it without you.

I also owe my success to my friends at home and abroad who have been there in spirit as I worked my way through this degree. Reggie, Sarah, Arash, Andrei, Ana, Stephen P., Abbie, Caitie, Kaitlyn, Stephen B., Megan, Mitch, Ryley, Curt, Karl, Ed, Kabir, Nathan, Kris, Kaylie, Dave, Amy, and Pearl.

Thanks always to my family, who has helped and supported me since the very beginning. Mom, you taught me the value of effective writing and communication and you're the best secret baking advice-giver I could ask for. Dad, you helped ignite my sense of wonder and curiosity in the mysteries of the universe. Holly, you've been my forever friend since my first moment in this world. Ba, you've always believed that I've had the smarts to do whatever I wanted, ever since I was very small. Thanks for the love and the support – I love you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Compact Stellar Remnants

Since the discovery of the first observed white dwarf 40 Eridani B, the degenerate remnants of stellar evolution have created mysteries that cannot be explained with conventional theories of physics. The development of stellar spectral classifications and plotting stellar populations on the Hertzsprung-Russell diagram (HR diagram; see Figure 1.1) demonstrates that the majority of stars follow a relatively well-defined temperature/luminosity correlation in a narrow part of the diagram. This part of the HR diagram is known as the main sequence, and is where stars spend the majority of their hydrogen-burning lifetimes. Examining the main sequence, it can be clearly seen that more massive main-sequence stars are hotter and more luminous than their less-massive counterparts. However, both 40 Eridani B and the white dwarf Sirius B seemed to contradict this idea (Holberg, 2005). Despite being small and faint, both stars were white-hot with surface temperatures of $\sim 12,000$ and $\sim 25,000$ K, respectively (Liebert et al., 2005). Even more unusual is the fact that white dwarfs are incredibly dense - observations of the Sirius system's binary motion showed that Sirius B had roughly the same mass as the Sun, yet luminosity and distance measurements showed that its radius was around 1% that of the Sun's - implying that Sirius B had a density roughly 10^6 times larger than the Sun! This observational evidence demonstrated that the relationships governing the structure of ordinary stars were inadequate for white dwarfs. The development of Fermi-Dirac statistics, with the realization that white dwarfs were supported against gravitational collapse by electron degeneracy pressure, was ultimately successful in explaining the nature of these objects (Chandrasekhar, 1931).

The discovery of the true nature of white dwarfs was followed by the 1934 prediction by Walter Baade and Fritz Zwicky of the neutron star, an even more compact stellar remnant (Baade & Zwicky, 1934). However, both the neutron star and the black hole would not be found until the 1960s. The 1967 discovery of the first neutron star by Jocelyn Bell Burnell and Antony Hewish, as well as the 1964 discovery of the black hole Cygnus X-1 (depicted in Figure 1.2) also made it clear that stellar remnants must be described with post-Newtonian physics (Bowyer et al., 1965; Hewish et al., 1968). This is easily established by examining their characteristics. For example, the mass and radius of a neutron star are generally taken to be ~ 1.4 solar masses ($M_\odot$) and ~ 10 km, respectively (Frank et al., 2002). The exact size of a neutron star of a given mass depends on the details of the equation of state for the neutron

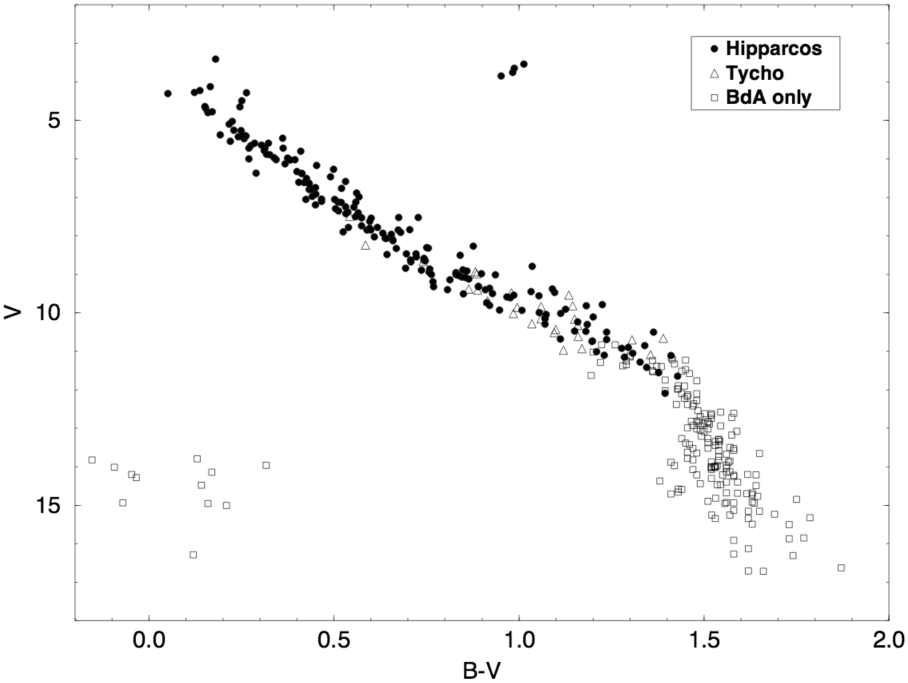

Figure 1.1: Hertzsprung-Russell diagram, also known as a colour-magnitude diagram (CMD) constructed for stars in the nearby Hyades cluster (distance: 46 pc), using photometry from the *Hipparcos*, Tycho, and BDA catalogues. The y-axis plots increasing brightness in the V filter corresponding to the luminosity of the source. The x-axis plots increasing red colour of objects, which corresponds to decreasing temperature. The middle of the plot is occupied by the main sequence of the Hyades cluster, where stars spend the majority of their hydrogen-burning lifetimes. As one moves from colder to hotter stars along the main sequence, the stellar luminosity increases (mass also increases, though it is not plotted here) along a tightly correlated curve. Along the bottom left of the plot are the white dwarfs, which exist outside the normal temperature-luminosity correlation for main sequence stars by being both very faint and very hot. This figure was taken from Perryman et al. (1998).

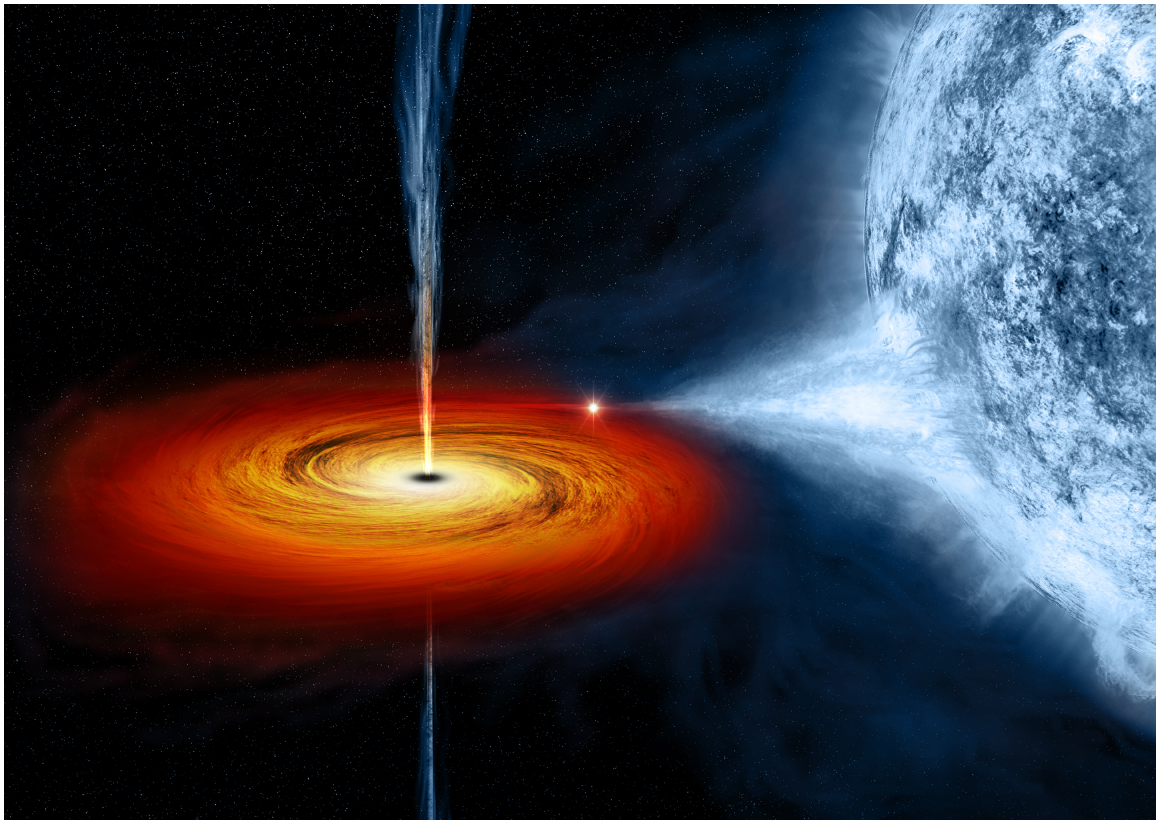

Figure 1.2: Artist's impression of the black hole X-ray binary Cygnus X-1. On the right, a blue supergiant (spectral class O9.7) companion is transferring mass via a stellar wind, as evidenced by the accretion stream in the middle, which is hotter at the point where it reaches the reddish-orange accretion disk on the left. The accretion disk encircles a $\sim 20 M_{\odot}$ black hole, and becomes steadily hotter at radii closer to where matter falls onto the black hole's "surface". This image also depicts radio jets, extended emissions of relativistically accelerated particles that are thought to remove angular momentum from the system (See Livio, 1997 and references therein). Image credit: NASA/CXC/M. Weiss.

star's interior, which are still unknown.

Attempting to use ordinary Newtonian mechanics to calculate a neutron star's escape velocity at the surface can demonstrate this theory's inadequacy in describing such an object.

$$\frac{m v_{\mathrm{esc}}^{2}}{R} = \frac{GMm}{R^{2}} \tag{1.1}$$

Solving for the escape velocity in Eqn 1.1 using the typical neutron star mass and radius values gives a value of roughly $v_{\mathrm{esc}} = 1.4 \times 10^{8}$ km s^{-1}, or about $0.455c$. Since the mechanical properties of particles (momentum, kinetic energy, etc.) begin to deviate relativistically from classical predictions at about $0.1c$, discussion of compact objects clearly requires relativistic theory.

The extreme nature of white dwarfs, neutron stars, and black holes make them exceptionally useful for testing predictions of modern relativistic theory, as they possess temperatures, pressures, densities, and magnetic field strengths that are difficult or impossible to replicate in terrestrial laboratories. Observing these objects in isolation is particularly challenging. Unlike main-sequence stars, white dwarfs and neutron stars do not fuse hydrogen as a source of energy. With a few exceptions, their only energy emitted is thermal radiation as they cool from their high birth temperatures. In the case of black holes, their gravitational pull is so strong that even light cannot escape, and they have no emission except perhaps for extremely weak blackbody radiation known as Hawking radiation (Lewin et al., 1997). The resolution to this observational complication lies in the existence of binary systems. The intense gravity of these compact objects permits them to accrete material away from their stellar companions when they exist in binaries. Instead of extracting energy from nuclear fusion (as ordinary main-sequence stars do), emission from a compact object binary system is powered by the gravitational infall of material onto the compact object. An order of magnitude estimate can show the relative efficiency of this process to stellar fusion. For hydrogen, approximately 0.7% of its mass is converted to helium, which implies that the energy release per unit mass is

$$\frac{\Delta E}{m} = 0.007c^2 \approx 6 \times 10^{14}\ \mathrm{J/kg}. \tag{1.2}$$

By contrast, consider the gravitational potential energy released as a mass m falls from a large distance onto a spherically symmetric body of mass M and radius R. In this case, the energy released per mass is

$$\frac{\Delta E}{m} = \frac{GM}{R}. \tag{1.3}$$

Using typical neutron star values (for example), one gets an energy release per mass of $\sim 2 \times 10^{16}$ J/kg, roughly a factor of 30 times more efficient than traditional nuclear fusion (Frank et al., 2002). This makes compact object accretion amongst the most efficient systems known at extracting energy.

An interesting consequence of conservation of angular momentum is that material that is transferred to the compact object from the companion cannot fall directly onto its surface, and instead must form an accretion disk around the compact object. As the material (usually hydrogen) falls onto the surface of the compact object, it loses a large amount of gravitational potential energy, converted to thermal energy by viscous and frictional forces (Lewin et al., 1997). This causes the accretion disk to have a high temperature, and an approximate energy scale for the emission from the disk can be derived by treating the disk approximately as a blackbody.

At the lower bound, if all the gravitational energy were radiated away as a blackbody, it would have a temperature defined by the Stefan-Boltzmann Law:

$$L = 4\pi R_\star^2 \sigma T_b^4 \tag{1.4}$$

thus

$$T_b = \left(\frac{L}{4\pi R_\star^2 \sigma}\right)^{\frac{1}{4}} \tag{1.5}$$

where σ is the Stefan-Boltzmann constant, and L and $R_\star$ are the luminosity and size of the accreting object, respectively.

The upper bound on the temperature of the accretor is given by assuming that all of the gravitational potential energy is converted to thermal energy, with the additional assumption that the accreted material consists of proton-electron pairs:

$$\frac{GM(m_p + m_e)}{R_\star} \approx \frac{GMm_p}{R_\star} = 2 \times \frac{3}{2} kT_{th}, \tag{1.6}$$

yielding

$$T_{th} = \frac{GMm_p}{3kR_\star}, \tag{1.7}$$

where m_p and m_e are the proton and electron masses, k is the Boltzmann constant, G is the gravitational constant, and M is the mass of the compact object. The bounds on the temperature are therefore

$$T_b \leq T_{\text{compact object}} \leq T_{th} \tag{1.8}$$

Taking the canonical mass (1.4 $M_\odot$) and radius (10 km) of a neutron star with a characteristic luminosity of $\sim 10^{36}$ erg s^{-1} gives the following:

$$6 \times 10^6\mathrm{K} \leq \mathrm{T_{NS}} \leq 10^{11}\mathrm{K} \tag{1.9}$$

$$0.5\mathrm{keV} \leq \mathrm{E_{photons,NS}} \leq 60\mathrm{MeV}. \tag{1.10}$$

Equivalently, for a white dwarf, a luminosity of $\sim 10^{33}$ erg s^{-1} and a radius of $\sim$ 7000 km gives the following temperatures and energies:

$$4 \times 10^4\mathrm{K} \leq \mathrm{T_{WD}} \leq 10^9\mathrm{K} \tag{1.11}$$

$$1\mathrm{eV} \leq \mathrm{E_{photons,WD}} \leq 100\mathrm{keV}. \tag{1.12}$$

From these calculations, we see that accreting white dwarfs should be optical to X-ray emitters, while neutron stars should be strong X-ray to γ-ray emitters. The calculations for a black hole yield roughly similar results as for a neutron star. This tells us that most binaries with a neutron star or black hole accretor should be X-ray binaries (XRBs).

Broadly, X-ray binaries tend to fall into two categories determined by the mass of the companion and the mode of accretion from the donor (Camenzind, 2007). Binaries with a low-mass donor, known as Low-Mass X-ray Binaries (LMXBs), tend to have a companion which fills its Roche lobe (a surface of gravitational equipotential within which material is bound to the companion) and transfers matter to the accretor through the L_1 Lagrange point between the two systems. Lagrange points are the locations in a two-body system where the gravitational force of the two objects on a test mass is equal to the centripetal force necessary to orbit them both at a constant, stable rate. The low luminosity of their companion makes LMXB systems dominated by their X-ray emission from the disk and (potentially) the compact object. Binaries with a high-mass donor are known as High-Mass X-ray Binaries (HMXBs). HMXBs usually accrete material through strong stellar winds driven by the companion, rather than Roche Lobe overflow. These systems tend to be optically bright because of the high luminosity of the O and B type stellar companions. Other forms of compact object binary exist, such as the Cataclysmic Variables (CVs), binaries consisting of a white dwarf with a main sequence donor, and ultracompact binaries, which consist of two compact objects in a close binary. These systems are often difficult to detect outside of the Milky Way because of their lower luminosities, with the exception of a class of white dwarf accretors known as supersoft X-ray sources (Kahabka, 2002).

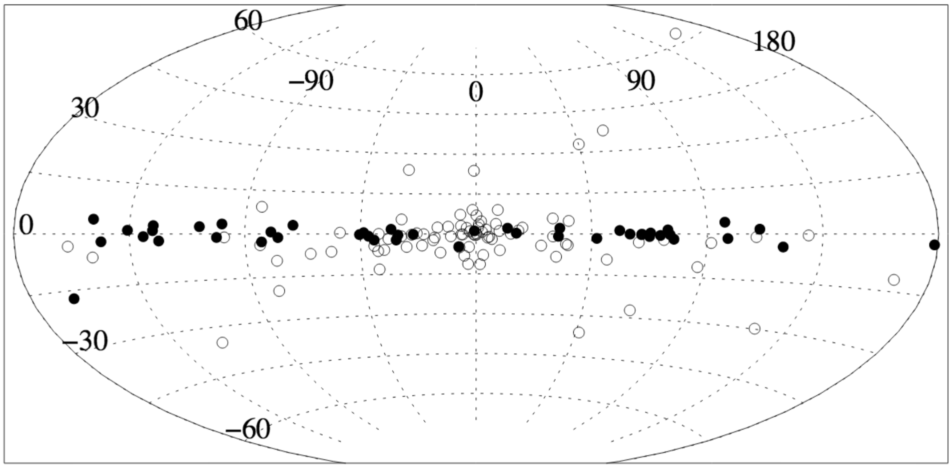

Figure 1.3: Distribution of XRBs located inside the Milky Way. This image is in Galactic coordinates, placing the Galactic Center at the middle of this figure. LMXBs are marked with white circles, while HMXBs are marked with black circles. Note that HMXBs and LMXBs are focused in the Galactic plane, while LMXBs in particular tend to be clustered about the Bulge. Figure taken from Grimm et al. (2002).

1.2 X-ray Binary Formation

One of the first observations made about XRBs was that they are not uniformly distributed throughout the Milky Way. An examination of the distribution of Milky Way's X-ray binaries shows that they are preferentially found in the Galactic plane rather than above or below it. Additionally, the HMXBs tend to cluster along lines of sight that correlate well to the Milky Way's spiral arms, regions of strong star formation (Grimm et al., 2002). Meanwhile, the abundance of LMXBs near the Galactic Bulge is indicative of the very high stellar densities found in the vicinity of the Galactic Centre (see Figure 1.3). The distribution of XRB sources reflects the complexity of the formation pathways of these sources. Initially, the fact that compact objects were found in binary systems was somewhat surprising. The progenitors of a black hole or neutron star are very high mass stars, and from simple energy considerations the expectation was that if half the mass (or more) of the primordial binary system was lost, the companion star would assume a hyperbolic orbit and the binary system would be unbound (Blaauw, 1961). In the case of higher-mass companions (where the mass ratio between the companion and the compact object progenitor is less extreme) it was evident that the survival of such systems was easier to justify, especially if mass-transfer had taken place from the progenitor to the companion (van den Heuvel & Heise, 1972; Tutukov & Yungelson, 1973). In the case of an LMXB, however, it is not clear how such a system could survive the compact object formation process, especially for systems where the orbit of the companion was inside the radius of the presumed progenitor star. In order to explain these systems, two key concepts were developed.

The first idea for explaining the survival of LMXB systems was the notion that the supernova explosions which form neutron stars and black holes are not symmetric. Instead, there is an asymmetry in the explosion that gives a "kick" to the compact object, giving it a non-trivial velocity relative to its progenitor. If such a kick is in a fortuitous direction ("tuned") then it can stabilize the binary against the loss of most of its mass and shrink the orbit (Flannery & van den Heuvel, 1975). Observations of the peculiar velocities of pulsars, rapidly rotating neutron stars that emit focused beams of radiation, provided evidence to suggest that such kicks occur, though it is not well-understood how the supernova explosion (or the initial core collapse) occurs asymmetrically (Lyne & Lorimer, 1994; Podsiadlowski et al., 2005).

A second necessary theoretical development was the concept of the "common envelope" (CE), a short phase in a binary system where a companion (often a compact one) will be engulfed by a giant star's envelope as it expands, and friction between the companion and the envelope will shrink the orbit even further. This phase ends with the ejection of the envelope, leading to the formation of a binary with a significantly reduced orbit, or potentially a stellar merger (Ivanova et al., 2013). Such an event is generally believed to be necessary for primordial XRB formation, given that the orbital distance of the companion star is usually inside of the radius of the compact object progenitor. However, investigations into common-envelope events are often limited by the computational complexity required to treat the binary system with a three-dimensional fluid dynamics (hydrodynamics) simulation. Like many astrophysical simulations, CE events require tracking the evolution of a binary system over many orders of magnitude in size, ranging from the ~ 10 km radius of a neutron star or smaller to the ≥ 1000 $R_\odot$ size of a giant star. Additionally, the time steps range from the short time scales (of order seconds) which capture the fine details of the CE event to the larger time scales of $\sim 10^3$ yrs for the duration of a CE event. The wide range of timescales involved requires that any individual CE simulation is likely to overlook important physics necessary to the process.

An additional complication is the necessity for CE events to explain the evolution of rare astrophysical systems aside from XRBs (Type Ia supernovae, gamma-ray bursters, and neutron star-neutron star binaries; Podsiadlowski, 2008). Common envelope events are relatively brief phases (of order 10^{2-3} yrs) in the lifetime of a binary system, making observing them on the sky a rare and difficult proposition. As a consequence, the majority of the observational data available to constrain simulations of binary formation is the existing population distribution on the sky of varied objects that are thought to require CE in their formation (Ivanova et al., 2013). However, such codes (known as population synthesis codes) must treat CE events in a simplified way in order to produce post-CE systems in a computationally tractable manner. This in turn can lead to an "overtuning" of binary population synthesis codes in order to reproduce observed populations, which is problematic since it is not clear whether alternate formation mechanisms for producing compact objects exist. The globular clusters of the Milky Way, discussed below, provide at least one alternate pathway for producing compact objects.

To illustrate the relative rarity of XRBs compared to main sequence stars, let us consider the number of XRBs that we might find in a stellar population. In this case, we will take a globular cluster (GC), as it is the most favourable to XRB formation (see below). For simplicity, we will ignore potential effects due to metallicity and the proximity of the globular cluster to its host galaxy. Consider a $10^5 M_\odot$ GC. First the number of compact object progenitors can be calculated from the initial mass function (IMF), which is defined as follows:

$$\frac{dN}{dm} \propto m^{-\alpha}. \tag{1.13}$$

This equation describes the number of stars N in a given population with mass m, using a power law with index α. Here, for a reasonable choice of IMF, α is taken to be 1.3 for $0.08M_{\odot} \leq m \leq 0.5M_{\odot}$ and 2.3 for $m \geq 0.5M_{\odot}$ (Kroupa, 2001).

Taking the minimum mass for hydrogen burning to be $0.08M_{\odot}$, the maximum mass of the IMF to be $100M_{\odot}$ and the minimum progenitor mass required to make a neutron star or black hole remnant to be $8M_{\odot}$, the fraction of stars forming black holes or neutron stars is given by:

$$f_{\rm rem} = \frac{\int_{8M_{\odot}}^{100M_{\odot}} m^{-2.3} dm}{\int_{0.08M_{\odot}}^{0.5M_{\odot}} m^{-1.3} dm + \int_{0.5M_{\odot}}^{100M_{\odot}} m^{-2.3} dm} \approx 0.01. \tag{1.14}$$

Next, consider the fraction of systems that are binaries. Such a value will scale with total cluster mass, but with a reasonable assumption that all binaries are found in the core, the fraction of binary systems (all binaries, not just those in compact objects) in the core of a globular cluster can be reasonably approximated by the following empirical relation:

$$f_{b,\rm core} = 0.13M_V + 1.07, \tag{1.15}$$

where M_V is the total absolute visual magnitude of the cluster (Leigh et al., 2011). Taking the mass-to-light ratio (in solar units) of the cluster to be ~ 2 for old globular clusters (see, for example, Strader et al., 2009), then the binary fraction in the core can be rewritten as:

$$f_{b,\rm core} = 0.13\left(4.77 - 2.5\log_{10}\left(\frac{10^5}{2}\right)\right) + 1.07 \sim 0.16. \tag{1.16}$$

The number of binary systems in the core is then given by $N_{bin,\rm core} = f_{b,\rm core} N_{\rm core}$. The number of total objects in the core, $N_{\rm core}$, can be approximated by assuming $M_{\rm core} = 0.1M_{\rm GC}$ and dividing it by $0.5M_{\odot}$, a reasonable average stellar mass for objects in the cluster. For a $10^5 M_{\odot}$ cluster, this gives $N_{\rm core} \sim 200,000$ objects, of which $N_{bin} = 0.16 \times 2 \times 10^4 \approx 3000$ are binaries.

Finally, multiplying the binary number by the compact object fraction (and assuming that binary fraction is a constant function of mass) gives

$$N_{XRB} = N_{\rm bin,core} * f_{\rm rem} \approx (3000)(0.01) = 30 \tag{1.17}$$

XRBs in a cluster of 10^5 solar masses. Although this is an overestimate (for example, it does not account for the possibility of systems consisting of two compact objects) it clearly illustrates the relative rarity of XRBs. Even in an environment comparatively well-suited to XRB formation, only 0.03% of the objects in a GC are in an XRB.

Despite the small numbers of XRBs throughout the Milky Way and its satellites, it has been known since the early days of X-ray astronomy (with the launch of the first X-ray satellite *Uhuru*) that many bright X-ray sources are located in the Milky Way's globular clusters (Clark, 1975). In fact, it was soon discovered that X-ray binaries are roughly a factor of 100 times more abundant (per unit mass) in globular clusters than they are in the Milky Way itself, and roughly 10% of the Milky Way's LMXBs are located in clusters rather than in the field (Verbunt & Lewin, 2006). Unlike X-ray sources in the Milky Way field, globular clusters do not host HMXBs. Globular clusters are devoid of gas, are described by only a few star formation epochs in their history, and have characteristic ages of $\geq$ 10 Gyr, factors which mean they no longer have any short-lived, high-mass stars. However, they do possess an abundance of compact object systems, some of which emit in X-rays, including numerous LMXBs, cataclysmic variables, and pulsars (possessing millisecond or longer spin periods) (Heinke, 2010). XRBs are often easier to study in globular clusters because cluster distances and ages are comparatively well-known, and many globular clusters are found at high Galactic latitudes, reducing the absorption of short-wavelength light due to interstellar dust and gas.

The overdensity of X-ray binaries in globular clusters was immediately attributed to their high stellar density, which permits X-ray binaries to form through dynamical interactions (rather than binary evolution, as for XRBs in the field; Verbunt & Lewin, 2006). Examples of these dynamical interactions include a three-body interaction where a compact object is exchanged with one member of an existing binary, tidal capture of a companion by a compact object, or a collision of a compact object (especially a neutron star) with a giant star. Examination of Milky Way and M31 globular clusters in the 80's and 90's, respectively, seemed to indicate a trend that clusters which contained more mass, possessed the highest metal abundances, and were the most compact tended to have the most X-ray binaries (Grindlay, 1987; Bellazzini et al., 1995). However, the number of X-ray binaries is generally thought to be best related to the stellar encounter rate Γ, a measure of the interaction rate per unit mass, which is typically parameterized in terms of a cluster's density ρ and velocity dispersion σ as follows:

$$\Gamma \propto \int \frac{\rho^2}{\sigma}\, dV. \tag{1.18}$$

Careful measurements of the stellar encounter rate for the Milky Way's globular clusters using surface brightness deprojection have revealed that there is a strong correlation between Γ and the number of LMXBs inside of a cluster, as shown in Figure 1.4 (Bahramian et al., 2013; Agar & Barmby, 2013). However, there is still considerable uncertainty on the exact strength of this correlation, due to small number statistics and uncertainty on how the internal structure of the cluster (beyond its global Γ) contributes to the XRB formation processes. For example, most globular clusters follow a luminosity profile that increases with decreasing radius from the core but flattens out within the central few parsecs. However, many globular clusters are core-collapsed, and exhibit luminosity profiles that increase all the way to the core (Harris, 1996). Core-collapse clusters exhibit a tendency (though not a well-defined statistical correlation) to possess fewer XRBs than other clusters with similar Γ values (Bahramian et al., 2013). It is not clear whether core-collapse clusters are ejecting XRBs more often than other clusters, or whether the presence of XRBs within clusters can prevent or reverse core collapse through the transfer of kinetic energy to stars in the core from tight binaries via three-body encounters.

The relative abundance of XRBs in globular clusters compared to the field led to the suggestion that many XRBs are formed inside of globular clusters and are transferred to a galaxy through dynamical ejection or the disruption of the globular cluster by the tidal field of the host galaxy (Grindlay, 1988; White et al., 2002). Investigation of this theory inside the Milky Way is difficult, since the extent of the interactions between the MW and its globular cluster system in earlier epochs is difficult to constrain based on present-day observations. Therefore, the best environment for testing whether galaxies receive their XRBs from globular clusters would be galaxies that generally lack globular clusters of their own. As discussed below in Section 1.4, the dwarf galaxies are suited for evaluating this theory for a variety of reasons.

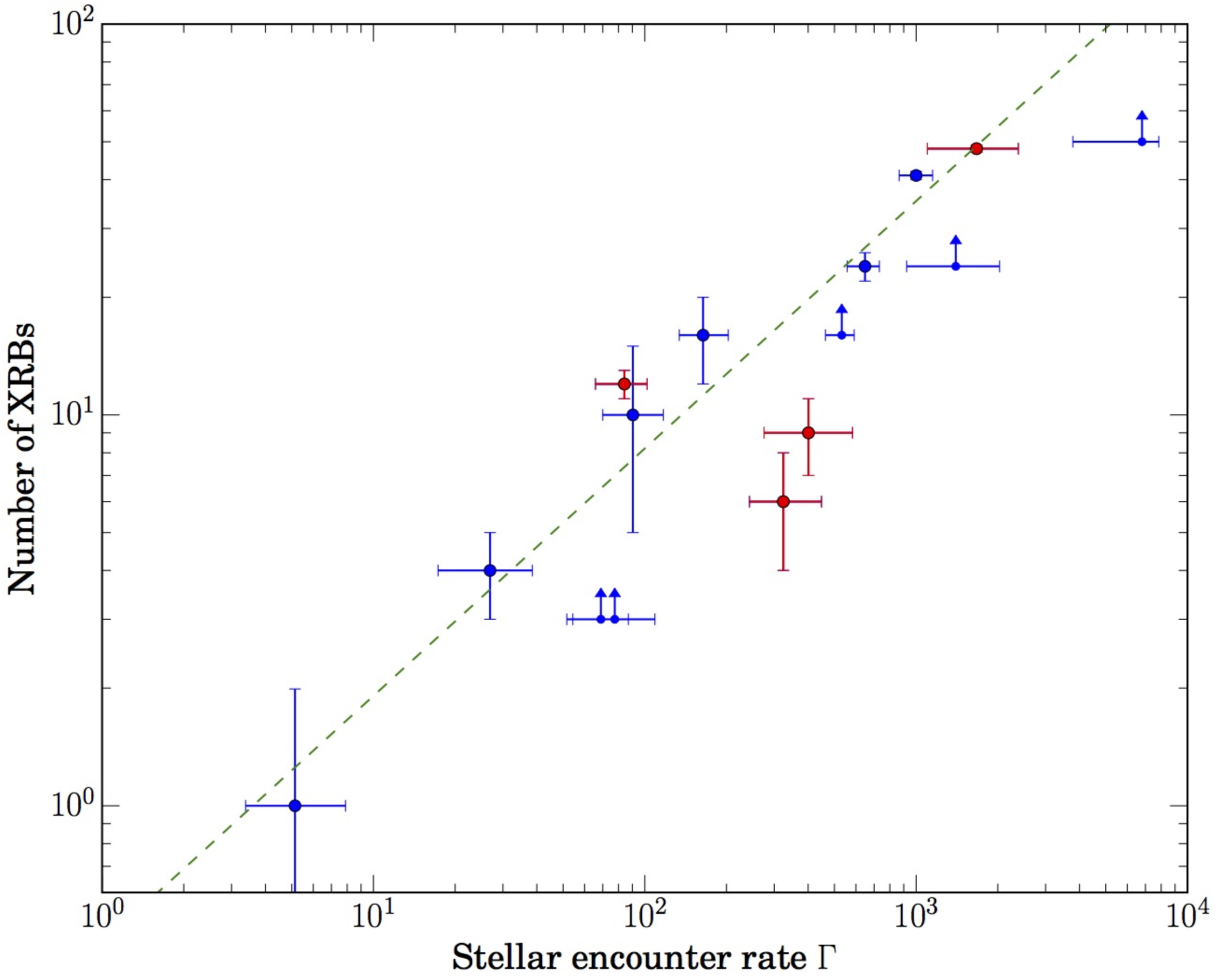

Figure 1.4: Number of XRBs contained in a globular cluster as a function of its stellar encounter rate. Core-collapsed clusters (see text) are in red, while other clusters are marked in blue. Figure adapted from Bahramian et al. (2013).

1.3 Galaxies and the Local Group

The Local Group is the gravitationally bound group of galaxies that includes our Milky Way, the Andromeda Galaxy (M31), the Triangulum Galaxy (M33), M32, and a host of satellite dwarf galaxies. Andromeda and the Milky Way are the two largest galaxies of this group and in fact a simple choice to model the Local Group is to consider it a binary galaxy group with galaxies clumped around the Milky Way and M31. M31 and the Milky Way account for $\sim 90\%$ of the Local Group's luminosity. These two galaxies provide the easiest pathway in the nearby universe for understanding the formation and evolution of massive spiral galaxies in a low density environment.

From our position inside the Milky Way, located in the Orion spur of the Sagittarius Arm, we have the ability to study the local structure and content of a spiral galaxy on the scale of individual stellar systems. This means that the Milky Way is crucially important as a cornerstone for understanding the properties of any astrophysical system that is too faint and/or too small to be seen in nearby galaxies. The tradeoff to this detailed view is that it is highly biased, as our position within the Milky Way's disk means that many lines of sight are heavily obscured by the presence of dust and gas in the disk. There exists a band on the sky called the "Zone of Avoidance", approximately 20 degrees thick and at all Galactic longitudes (in total, around 20% of the sky) where Galactic dust blocks optical observations (Jarrett et al., 2000; Kraan-Korteweg & Lahav, 2000). The Zone of Avoidance primarily has the observational effect of creating a significant gap in observing extragalactic sources. An additional effect is that the contents of the other side of the Milky Way are difficult to study. This means that structural properties of the Milky Way, such as spiral arm number, location, pitch, origin, and shape, require careful modelling and use of appropriate tracers to constrain constrain them. Moreover, integrated luminosity functions of populations, the rotation curve, and the Milky Way's total luminosity are not straightforward to measure. Nevertheless, studies of the Milky Way indicate that it is a relatively ordinary spiral galaxy of type S(B)bc I-II (Reshetnikov, 2000). This Hubble classification tells us that the Milky Way is a spiral galaxy (S) with a bar (B), a modestly sized bulge (bc) that is relatively bright (I-II).

The Andromeda Galaxy (M31) provides an opportunity to study a nearby spiral galaxy that is approximately analogous to the Milky Way. Indeed, the realization that M31 is an extragalactic structure in the Universe (Hubble, 1929) could be thought of as a cornerstone of extragalactic astronomy. Unlike the Milky Way, we can view M31's morphology and structure at a more favourable inclination of 78° relative to the normal, and being outside this galaxy allows us to observe all of its components. This means that M31 is the second stepping stone (after the Milky Way itself) in understanding the formation and evolution of massive spiral galaxies. The Milky Way and M31 share many similarities. They have similar bulge sizes, disk sizes, disk luminosities, and H I gas content (see van den Bergh 2000 and references therein). They also have relatively similar Hubble types, with M31 being a slightly more bulge-dominated galaxy with a Hubble type of Sb. Measurements of both M31 and the Milky Way point to M31 being more massive and luminous, and having a more luminous bulge. Despite this, M31 has a lower star formation rate than the Milky Way, placing these two galaxies in a distinct evolutionary position that is separate from all other galaxies in the Local Group – they are neither as active in star formation as the Magellanic clouds, but they are not "red and dead" like the many dwarf galaxies orbiting the major galaxies of the Local Group (Robitaille

& Whitney, 2010; Rahmani et al., 2016). M31 has a globular cluster system which is similar to the Milky Way's in many respects, though there are significantly more globular clusters in M31 (Barmby et al., 2000). However, recent studies have shown no evidence for obvious bimodality in M31's globular cluster system, distinct from the Milky Way's GC system (Caldwell et al., 2011). This implies that the formation history of M31, or at least its globular cluster system, is more complex than the Milky Way's.

The many dwarf galaxies of the Local Group represent a fascinating intermediate cosmological structure between the smaller and denser globular clusters and the much larger giant galaxies. Many of these objects were discovered only recently. They are the most abundant galaxy type in the universe (of the ~ 100 identified galaxies in the Local Group, at least ~ 90 can be considered dwarf galaxies depending on the demarcation criterion), but are also extremely challenging to study observationally. Their star counts are intermediate between the $\sim 10^6$ stars in a globular cluster and the 10^{11-12} stars of a giant galaxy, possessing at most a few 10^9 stars. Their characteristic sizes are a half-light radius of roughly $100 - 1000$ pc and their stellar masses are roughly $10^{6-8} M_\odot$ (McConnachie, 2012). Half-light radius is the radius within which half the light coming from a galaxy is contained, and is a useful shorthand for determining an object's spatial extent. Dwarf galaxies are larger than globular clusters both in terms of size and mass, but they have central densities roughly 100 times lower than a globular cluster. This means that dwarf galaxies are very faint and diffuse with low central surface brightnesses. Therefore, unlike globular clusters and giant galaxies, both of which have been detected since the 17th century or earlier, true dwarf galaxies (aside from the Small and Large Magellanic Clouds) only began to be discovered in the 20th century. Since the advent of sensitive, high angular resolution telescopes like the *Hubble Space Telescope* and the automation of searches for new dwarf galaxies from surveys, there has been an explosion in the number of discovered dwarf galaxies (McConnachie, 2012). The first dwarf galaxy discovered within the Milky Way system was the Sculptor Dwarf Spheroidal Galaxy. Its discoverer, Harlow Shapley, described it as being either like a galaxy-sized globular cluster or a nearby spheroidal galaxy with exceptionally low surface brightness, highlighting the complications in classifying these objects (Shapley, 1938).

As with giant galaxies, dwarf galaxies have a number of different subtypes based on their morphology and observed properties. The four principal kinds of dwarf galaxy, first outlined in van den Bergh (1959), are the dwarf spheroidals (dSph), dwarf ellipticals (dE), dwarf spirals (dS), and dwarf irregulars (dIrrs). Dwarf spheroidals and ellipticals structurally resemble giant ellipticals, although they have very low surface brightnesses and masses and sizes in the range defined above. The distinction between dEs and dSphs is a relatively arbitrary luminosity and absolute magnitude cut: those with $L \geq 3 \times 10^7 L_\odot$ or $M_V \leq -14$ are the dEs, while fainter than this limit are the dSphs (Sparke & Gallagher, 2007). The majority of the galaxies in the Local Group are dwarf spheroidals (McConnachie, 2012). Dwarf spiral galaxies are those in the above mass/size/star count range that exhibit spiral structure, and are comparatively rare - there are no dwarf spirals in the Local Group (van den Bergh, 1959). Finally, dwarf irregular galaxies are those that lack identifiable symmetry, and are defined as irregular galaxies with $L < 10^8 L_\odot$. They tend to be comparatively brighter than other dwarf galaxies, because they typically possess more active star formation (Sparke & Gallagher, 2007).

Initially, dwarf galaxies were thought to have similar structure and evolution to that of globular clusters, albeit with much lower surface brightnesses. Examining structural properties of

`C1_plots/tolstoy09.jpg`

Figure 1.5: Plot of structural properties for a variety of different galaxy types. The x-axis plots the absolute visual magnitude of objects decreasing to the right, (a measure of total luminosity), while the y-axis indicates the half-light radius of objects. Galactic globular clusters are marked as gray dots. In addition to a number of individual giant galaxies (marked with blue triangles), typical values for giant elliptical galaxies and dwarf irregular galaxies are marked with red and dashed gray ellipses, respectively. Classical Milky Way Dwarf galaxies within the Local Group are marked with pentagons, while the more-recently discovered ultra-faint dwarf galaxies are marked with yellow stars. Finally, the ultra-compact dwarf galaxies (discussed in the text, below) of the Virgo and Fornax clusters are marked with purple crosses. Across the diagram, a clear trend is noticeable. Dwarf galaxies, especially those that possess gas or are ultra-faint, tend to exist in a similar luminosity regime to that of the Galactic globular clusters, but are spatially much larger and more extended. Additionally, dwarf galaxies often possess comparable spatial sizes to ordinary elliptical galaxies, but they are much fainter. This figure was taken from Tolstoy et al. (2009). Copyright permission to reproduce this figure was not available, therefore we invite the reader to view it at `https://www.annualreviews.org/doi/pdf/10.1146/annurev-astro-082708-101650` (Figure 1).

dwarf galaxies (see Figure 1.5) shows similarities both with ordinary galaxies and with globular clusters. However, photometric studies of dwarf galaxies (especially in the Fornax Dwarf Galaxy) revealed the presence of carbon stars, which suggested multiple stellar populations and enhanced metallicity as compared with the metal-poor, (canonically) single-population globular clusters (Aaronson & Mould, 1980). In addition, measurements of the mass-to-light ratios of dwarf galaxies showed that many possessed mass-to-light ratios at least an order of magnitude larger than that of the globular clusters (Mateo, 1998). These exceptionally high mass-to-light ratios were used to infer the presence of dark matter halos in these galaxies, which globular clusters do not possess (Lane et al., 2010).

The observational distinction between giant galaxies and dwarf galaxies is similarly complicated. Initially, dwarf galaxies were thought to be entirely distinct from giant galaxies. Measurements of central surface brightness against luminosity for elliptical galaxies made early in the history of the field appeared to show two separate correlations, one for dwarf galaxies and one for larger galaxies, as shown in Figure 1.6 (Kormendy, 1985). These correlations suggest that dwarf galaxies form via a distinct mechanism from giant galaxies. However, studies conducted more recently instead imply that (at least in the case of elliptical galaxies) there is a continuous correlation between galaxies of varying sizes, suggesting that dwarf and giant galaxies may share more common origins (see Figure 1.7 Graham & Guzmán 2003).

The abundance of dark matter in dwarf galaxies makes them the most dark-matter dominated objects known, and reveals another observational challenge. The standard Λ Cold Dark Matter (ΛCDM) model of cosmology predicts that dark matter should cluster into halos of different sizes, with smaller halos being much more abundant than larger ones. However, cosmological simulations are only able to recreate the distribution of larger halos (i.e., that of ordinary galaxies) but predict far more small halos (such as those in dwarf galaxies) than are observed in the Local Group. Two possibilities for this so-called "missing satellite" problem exist: either small dark matter halos do not form stars efficiently, or the otherwise robust predictions of the ΛCDM model are incorrect (Klypin et al., 1999). An intriguing solution to this problem was presented by the discovery of a new subclass of dwarf galaxies, known as the ultra-faint dwarf galaxies (Willman, 2010). These galaxies possess total luminosities comparable to globular clusters, a spatial extent similar to classical dwarf galaxies, and velocity distributions that imply they are well over 99% dark-matter dominated. These systems are also tremendously metal-poor, implying that they are useful tools for near-field cosmology that permit the study of the Universe through structures that formed early in the history of the Universe (Bullock, 2010). In any case, the discovery of these ultra-faint dwarf galaxies provides empirical evidence that the "missing" satellites likely exist, but at this time are too faint and diffuse to have been observed with current search strategies.

Our picture of dwarf galaxies is complicated considerably by a number of observational constraints. The low surface brightness of dwarf galaxies makes them difficult to detect outside the Local Group. As a consequence, much of the theory of dwarf galaxy formation and evolution assumes that the dwarf galaxies we can observe are a representative sample of dwarf galaxies throughout the universe. This assumption is rather limiting, as even the dwarf galaxies associated with M31 appear to have significant differences from the Milky Way's dwarf galaxy population (McConnachie & Irwin, 2006). A number of intriguing dwarf galaxy subclasses, including the blue compact dwarf galaxies (which appear to be active sites of star formation) and the ultracompact dwarf galaxies (which have sizes comparable to globular clusters but

C1_plots/correlations.jpg

Figure 1.6: Plots of elliptical galaxy structural parameters as a function of luminosity (absolute magnitude). The top panel plots central surface brightness, while the bottom plot shows core radius (the radius at which the surface brightness of an object drops to half of its maximum value). Large circles are ordinary elliptical galaxies, large squares are dwarf ellipticals, crosses are irregular galaxies, and small dots are globular clusters. Note that there appears to be a separate correlation for giant and dwarf ellipticals. More recent work, such as Figure 1.7 and works citing that source paper, shows a more continuous correlation that joins elliptical galaxies of all masses. Figure taken from Kormendy (1985). Copyright permission to reproduce this figure was not available, therefore we invite the reader to view it at `http://adsabs.harvard.edu/doi/10.1086/163350` (Figure 3).

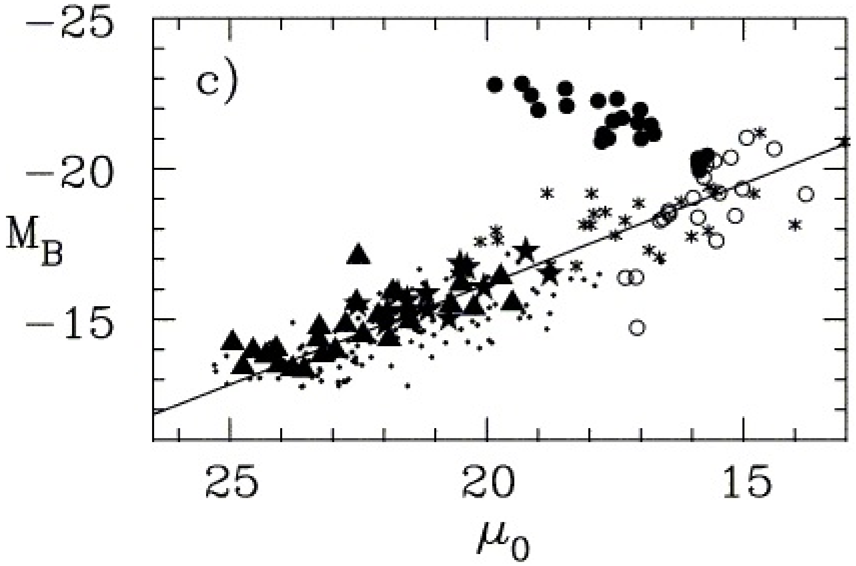

Figure 1.7: Plot of elliptical galaxy absolute magnitude as a function of central surface brightness. Objects with larger absolute magnitude are located at the top, while objects with a larger central surface brightness are located to the right. Dots, triangles, and stars represent dwarf elliptical galaxies taken from various studies, asterisks represent "intermediate" elliptical galaxies, and circles (filled in and open) represent giant elliptical galaxies from various studies. Figure adapted from Graham & Guzmán (2003).

are spectrally identified as dwarf galaxies) are absent from the galaxy population of the Local Group. These galaxies likely have different formation and evolution pathways from the Local Group's dwarf galaxies. Ultracompact dwarfs, for example, are thought to be the cores of dwarf elliptical galaxies that have been tidally stripped by interacting in dense galaxy clusters (Kazantzidis et al., 2004). It is also worth noting that observations of the Local Group's dwarf galaxies suggest that they possess a range of star formation histories (Weisz et al., 2014).

Another complication in considering dwarf galaxy formation and evolution is the fact that many dwarf galaxies seem to interact with their host galaxies. The most striking example of this is the Sagittarius Dwarf Galaxy, discovered only in 1994, which is in the process of being destroyed and absorbed by the Milky Way itself (Ibata et al., 1994). In fact, it is generally believed that new, young dwarf galaxies can be formed as a consequence of interactions between larger galaxies. These tidal dwarf galaxies are distinct from the primordial dwarf galaxies in that they are expected to lack dark matter when compared with the primordial dwarf galaxies, which form in dark matter halos early in the Universe (Kroupa et al., 2010).

Despite the complexity and observational challenges, dwarf galaxies (especially primordial dwarf galaxies) provide an interesting test environment to attempt to link various theories for the formation and evolution of ordinary galaxies with those of globular clusters. In particular, they are useful for examining theories of X-ray binary formation, discussed below.

1.4 X-ray Binaries in the Local Group

Beginning with the detection of Sco X-1 by early rocket-based experiments, the Milky Way's X-ray binary population remains, in many ways, the easiest to study (Giacconi et al., 1962). The modern understanding of the XRBs in the MW has been shaped by three approaches to characterizing its XRB population. The first of these is the development of a high spatial resolution all-sky survey enabled by the *ROSAT* telescope, which surveyed 95% of the sky in the 0.1–2.4 keV energy range with superior sensitivity to earlier X-ray telescopes (Voges et al., 1999). All-sky surveys enable us to obtain a picture of the Milky Way's X-ray binary population all at once in one energy range. This allows us to understand the Milky Way's XRB population as it would look to an external observer, similar to how we can image nearby galaxies and construct their luminosity functions. The massive angular size of the Milky Way means that only an all-sky survey can create this understanding.

A second approach that has greatly enhanced our understanding is through missions focused on the time domain. This includes observations taken with high timing resolution, such as the *Rossi X-ray Timing Explorer* (*RXTE*), as well as observatories which are able to provide information about and perform rapid observations of transient phenomena, such as the *Swift X-ray Telescope* (Bradt et al., 1993; Gehrels et al., 2004). Studies of transient phenomena, such as outbursts, pulsations, and eclipses, have given us clues into understanding the physical origin of these phenomena. These phenomena also let us understand the life-cycle and evolution of XRBs as a whole throughout the Milky Way. We understand the Milky Way's XRB population to be variable, where many sources have long periods of quiescence that may limit their visibility.

Finally, the Milky Way's XRB population has been better understood through the advent of high-angular resolution X-ray telescopes, including the *Chandra X-ray Observatory*, *XMM-*

Newton, and the *Nuclear Spectroscopic Telescope Array*, which provide observations across the soft and hard X-ray bands (Weisskopf et al., 2000; Jansen et al., 2001; Harrison et al., 2013). These telescopes permit us to study sources in the Milky Way both photometrically and spectroscopically, even in crowded areas, such as the Galactic Center of the Milky Way. Moreover, high-angular resolution (especially the sub-arcsecond resolution of *Chandra*) observations enable identification of counterparts at longer wavelengths because of accurate photometry. In many cases, Milky Way X-ray sources can be matched to unique optical counterparts, which permit us to constrain detailed parameters of individual systems, including distance. The sensitivity of telescopes such as *Chandra* also allows us to study the more challenging-to-observe quiescent population of XRBs, which are believed to comprise the majority of XRBs in the Milky Way at any particular time.

In contrast with the optical emission of the Milky Way, which is dominated by emission from ordinary stars, the X-ray binary population of the Milky Way, which comprises the bulk of the brightest X-ray sources in the Galaxy, comprises only a few hundred objects in total. As discussed in Section 1.2, the density of X-ray binaries along different lines of sight is a function of Galactic structure. In brief, HMXBs are concentrated in the disk, and are generally believed to be associated with recent star formation in the Milky Way's spiral arms, while LMXBs tend can be found inside and outside of the disk, though they are most abundant in the direction of the Galactic Center and in the globular clusters of the Milky Way. The Milky Way's cumulative X-ray flux distribution – that is, the number of sources N above a particular flux S, has a number of interesting properties that were revealed in a study using the *RXTE* all-sky survey (Grimm et al., 2002). This function is interesting in two ways – firstly, at the bright end it is dominated by contributions from only a handful (< 10) of XRBs. This, combined with the fact that individual XRBs can show variability between bright states and quiescence on timescales of days to years, means that the X-ray luminosity function of the Milky Way and other nearby galaxies can be variable on relatively short timescales. Additionally, the X-ray flux function appears to be dominated by contributions from LMXBs, which contribute $\sim$90% of the Milky Way's XRB luminosity compared to HMXBs. However, obscuration along certain lines of sight, as well as uncertain distance estimates to individual sources, means there is considerably more to explore in the Milky Way's XRB population and its relation to the Milky Way's formation and evolution.

As the nearest giant neighbour galaxy, it is unsurprising that M31 was the first galaxy where a population of LMXBs was discovered (Fabbiano, 2010), though it was not until the launch of the *Einstein* X-ray satellite that M31 appeared as a resolved collection of sources (van Speybroeck et al., 1979). X-ray sources in M31 have a lower relative uncertainty in their distance measurements compared to those in the Milky Way, because the distance to M31 is known with lower relative uncertainty compared with individual objects in the Milky Way's disk. In the years prior to the launch of higher-resolution X-ray platforms, studies of M31's X-ray population tended to focus on the properties of the galaxy as a whole, because lower resolution observations make it more difficult to associate X-ray sources to particular counterparts. This makes removal of background and foreground sources more challenging. *HST* has also helped deepen understanding of the X-ray source population of M31, because it allows X-ray sources to be associated with particular star clusters, and it can identify contaminating background galaxies and active galactic nuclei (AGN) relatively effectively. However, separation of different kinds of X-ray sources within M31 itself, such as identifying LMXBs or HMXBs, is still difficult at M31's distance.

The most complete (by area) high angular resolution survey of the majority of M31 was conducted with *XMM-Newton*. The resulting catalogue included some 1900 sources to a flux limit of a few 10^{34} erg s^{-1}, though the majority of these sources still lacked classification even with cross-matching with other X-ray, optical, and infrared catalogues (Stiele et al., 2011). *Chandra* observations of M31 have often focused on its bulge and globular cluster systems, since its field of view is small ($\sim 17'$) compared to the large angular size ($\sim 3^\circ \times 1^\circ$) of M31. The deepest study of M31, which combined all available archival *Chandra* data across the bulge and two fields in the disk, found that M31's X-ray luminosity function is a two-component power-law with a break at a few 10^{37} erg/s (Vulic et al., 2016). Additionally, the XLF for the disk versus the bulge appear distinct from other galaxies, with an XLF that drops off much faster in the disk than in the bulge. Bright HMXBs have yet to be confirmed in M31, confirming that star formation in the disk is weak.

Multiple studies have investigated the higher energy X-ray ($>$ 10 keV, often referred to as hard X-rays) population of M31 using *NuSTAR*. Amongst other results, these observations have helped separate the neutron star and black hole LMXBs in a number of M31 globular clusters (Maccarone et al., 2016). In hard X-rays the luminosity function appears to be dominated by contributions from LMXBs in globular clusters (Vulic et al., 2018). Finally, studies of the bulge, where the majority of *Chandra* observations have been concentrated, have also been conducted. All of the *NuSTAR* sources detected in the central region of M31 have counterparts in the previous *XMM-Newton* catalogue, though *NuSTAR* observations were able to identify several new X-ray binary candidates using additional spectral information available above 8 keV (Stiele & Kong, 2018). Comparisons between the luminosity function of M31 and the Milky Way tend to give a similar shape, though it was found that the slope of the luminosity functions differed more than by what one would expect from simply scaling to stellar mass, which would be expected in the case of an X-ray luminosity function dominated by LMXBs (Grimm et al., 2002).

Compared with the major galaxies of the Local Group, most of the satellite dwarf galaxies of the Milky Way have been studied relatively infrequently through X-ray observations. The reasons for using dwarf galaxies to test theories of X-ray binary formation are not immediately obvious. Their usefulness is best illustrated by a particular challenge to theory, namely the presence of persistently bright low-mass X-ray binaries (LMXBs) in old stellar populations. Theoretical models of XRB formation are difficult to constrain owing to considerable uncertainty about the details of the mechanisms (common envelope evolution, supernova explosions, etc.) involved, but it is expected that XRB production in a population should peak within $\sim 10^9$ years after the peak of star formation (White & Ghosh, 1998). This is problematic, as the age of LMXBs (which should be of order 10^{10} years) would exceed the mass transfer lifetime of their donor by a at least an order of magnitude.

As an example, consider accretion from a 1 $M_\odot$ donor onto a 1.4 $M_\odot$ neutron star. Making the simplifying assumptions that accretion is radiatively efficient and that 100% of the gravitational potential energy released by infalling material is converted to the outgoing X-ray luminosity, we can write the accretion luminosity as:

$$L_{acc} = \frac{G M_{\rm NS} \dot{M}}{R_{\rm NS}}, \tag{1.19}$$

where $M_{\rm NS}$ is the $\sim 1.4 M_\odot$ mass of the neutron star, $R_{\rm NS}$ is its $\sim$ 10 km radius, and $\dot{M}$ is the mass transfer rate between the donor and the neutron star (Frank et al., 2002). A typical X-ray luminosity of $\sim 10^{36}$ erg/s permits one to solve for the mass transfer rate, as follows:

$$\dot{M} = \frac{L_{acc} R_{\rm NS}}{G M_{\rm NS}} = \frac{(10^{38}\,{\rm erg/s})(10\,{\rm km})}{G 1.4 M_\odot} \approx 9 \times 10^{-9} M_\odot/{\rm yr} \tag{1.20}$$

Therefore, steady-state accretion would dictate that the 1 $M_\odot$ companion would be completely consumed by the neutron star within $1M_\odot/9\times10^{-9} M_\odot/{\rm yr} = 10^8$yr. Note that if accretion power is inefficient, the mass transfer rate needs to be even higher to sustain such high accretion luminosities.

Clearly, the difference between the ages of LMXBs in old populations and their expected accretion lifetimes is a factor of 10^2, creating a serious challenge to LMXB theory. Two main ideas have been proposed to resolve this problem, and we will consider each of them in turn.

The relative overabundance of XRBs in globular clusters (GCs), and the fact that they appear to be capable of producing new XRBs relatively recently in their evolutionary history, led to the suggestion that globular clusters may be the site for some (or all) LMXB production (Grindlay, 1988; White et al., 2002). Under this hypothesis, globular clusters efficiently produce LMXBs through dynamical mechanisms, and they are transferred to the host galaxy (e.g., the Milky Way) either by dynamical ejection or tidal effects. In the former case, the globular clusters would persist, but in the latter, the cluster may be disrupted entirely as it interacts with the host. However, there are a number of observational complications to the hypothesis that all XRBs come from GCs. The first is the results of studies that attempt to correlate integrated LMXB luminosity from a galaxy with galaxy parameters. Although it appears that the specific frequency of globular clusters (number of GCs per host galaxy luminosity) plays a role in LMXB abundance in galaxies, the strongest predictor of LMXB abundance within a galaxy seems to be stellar luminosity (and consequently, galaxy stellar mass) (Kim & Fabbiano, 2004; Gilfanov, 2004). A second challenge to the hypothesis is the observed distribution of LMXBs in other galaxies. Field LMXBs tend to be concentrated more towards the centers of their host galaxies than globular cluster LMXBs (Kundu et al., 2007; Kim et al., 2009). If GCs were ejecting LMXBs into the field of a galaxy, we would instead expect that GCs and Milky Way LMXBs should have roughly the same spatial distribution relative to the galaxy's center.

A second possibility is that the bright LMXBs are not actually persistently bright. This requires that these systems are transient and can change to periods of reduced accretion (lower mass transfer and consequently lower luminosity), known as quiescence (Piro & Bildsten, 2002). Under this theory, the LMXBs that are persistently bright over the existing $\sim$ 50 years of X-ray observations have long, unobserved periods of quiescence. A strength of this theory is the existence of a mechanism that permits LMXBs to switch between outburst and quiescence. This mechanism, known as the disc instability model, describes outbursts in LMXBs in terms of hydrogen ionization instability occurring in the accretion disk (Lasota, 2001). When the accretion disk around an LMXB is sufficiently heated by steady-state mass transfer, the disc becomes ionized and rapidly increases its viscosity, triggering an outburst and creating the observed bright LMXB behaviour (Done et al., 2007). A second advantage of this theory is that there is considerable observational evidence for LMXBs exhibiting transient behaviour both within the Milky Way and in other galaxies (Bahramian et al., 2014; Fragos et al., 2009). A notable requirement of this theory is that these persistently bright sources should have very long quiescent phases, well in excess of 75% of their accretion cycle (Piro & Bildsten, 2002). This would imply that for every apparently bright LMXB there should be a large number of quiescent LMXBs (qLMXBs). However, the numbers of qLMXBs existing in stellar populations are not well constrained because their faintness makes them difficult to detect.

There are a few mitigating factors for these theories of persistently bright LMXB systems. It is possible that some of these systems are ultracompact (a NS/BH accreting from a white dwarf), and these binaries are expected to evolve to a configuration where mass transfer happens much later after the peak of star formation occurs (Bildsten & Deloye, 2004). Additionally, it is possible that some of these systems were initially intermediate mass X-ray binaries (IMXBs) that evolved from an initial donor mass of $1 - 8M_{\odot}$ to a much smaller mass in the present era (Podsiadlowski et al., 2002). It is important to note that these systems are expected to be comparatively rare based on the initial mass function (IMF, see Section 1.2) of the stellar population (similar to the arguments made for the rarity of XRBs in Section 1.2). An additional consideration is that some persistently bright systems may have had no mass transfer early in their histories (and hence, would not have been an X-ray binary at all) until the donor was sufficiently evolved on the subgiant branch to fill its Roche Lobe. This is a likely explanation for systems that seem to exhibit no quiescence, such as Sco X-1 (Revnivtsev et al., 2011).

What makes dwarf galaxies ideal for evaluating these theories? To evaluate these theories, one would prefer a relatively isolated stellar population that lacks globular clusters of its own. Additionally, this stellar population should be old and dominated by a single epoch of star formation early in its history. Finally, this stellar population should be relatively diffuse in comparison to the globular clusters, so that dynamical interactions are largely unimportant.

Dwarf galaxies can host globular clusters of their own. An excellent example is the Sagittarius Dwarf Elliptical Galaxy, which has at least two globular clusters (M54 and Palomar 12) associated with it that were originally thought to be associated with the Milky Way instead (Cohen, 2004; Siegel et al., 2007; Sbordone et al., 2007). Dwarf galaxies can also, as discussed above, can be tidally disrupted by interacting with their host galaxy. However, most dwarf galaxies bound to the Milky Way are much further away from the Galactic centre than its globular clusters, typically above 40-50 kpc, and they mostly lack globular clusters of their own (McConnachie, 2012). Additionally, though some dwarf galaxies show evidence of multiple stellar populations, other dwarf elliptical and dwarf spheroidal galaxies bound to the Milky Way show evidence of only a burst of star formation early in the galaxy's history (Maccarone et al., 2005). Finally, the central stellar densities of these dwarf galaxies are roughly two orders of magnitude lower than those of globular clusters (Harris, 1996; McConnachie, 2012). Since they are relatively close to the Milky Way (compared with other galaxies, especially outside of the Local Group), transient LMXBs that are in their quiescent phase ($L_X 10^{31-33}$ erg/s) can potentially be discovered if observations are sufficiently deep.

Aside from answering questions about their own evolution, LMXBs in dwarf galaxies can help constrain properties of dwarf galaxies. As discussed above, LMXBs can be born with considerable "kick" velocities due to the recoil from the supernovas explosion that create them, and in many cases these velocities exceed the local stellar velocity dispersion inside of the galaxy. If a dwarf galaxy consisted only of the baryonic, luminous matter that can be observed, then we would expect that dwarf galaxies would be unable to hold onto LMXBs because of the shallowness of their potential well. However, if a dwarf galaxy possesses dark matter halos, then it is possible for it to retain its LMXBs. It is also possible that LMXBs are born with a distribution of kick velocities that includes lower velocities (< 20 km s^{-1}) permitting them to be retained by a dwarf galaxy's gravitational potential (Podsiadlowski et al., 2005). This was applied to the case of the Sculptor Dwarf Spheroidal Galaxy to independently constrain the mass of its dark matter halo (Dehnen & King, 2006).

X-ray studies conducted in the last two decades have been focused on two properties of dwarf galaxies. The first is that they are candidates for hosting intermediate mass black holes (IMBH) in their centers. IMBHs are a proposed class of black holes that lie between the stellar mass black holes and the supermassive black holes at the center of galaxies. The theoretical basis for this is that extrapolating the scaling relation between the mass of a galaxy's bulge and the mass of its central black hole for ordinary galaxies down to dwarf galaxy masses gives a mass of roughly $10^{4-5} M_\odot$, the mass range for IMBHs. X-ray searches of dwarf galaxies have found potential evidence for IMBHs at the centers of the Ursa Minor Dwarf and Draco galaxies, but this evidence is only tentative and lacks dynamical confirmation (Nucita et al., 2013; Sonbas et al., 2015). The presence of ultraluminous X-ray sources (ULXs; X-ray sources with luminosities $\geq 10^{39}$ erg s^{-1} that are considered potential IMBH candidates, see Swartz et al. 2004) has also been considered from a theoretical standpoint, though no ULX has been detected in a dwarf galaxy (Swartz et al., 2008).

The second property of nearby dwarf galaxies that has led to X-ray observations is that they are dark-matter dominated. Their extreme mass-to-light ratios implies that ordinary astrophysical (i.e., stellar) emission is comparatively weak. Attempts to determine the most likely dark matter particle involve searches for potential annihilation or decay products, which includes photons in the X-ray to γ-ray bands (Bergström, 2012). A number of searches for dark-matter signals in dwarf galaxies have been conducted. Studies in the γ-ray band have generally focused on attempting to constrain the strength of emission using a parameterized weakly interacting massive particle (WIMP) model (Abramowski et al., 2014; Bonnivard et al., 2015). By contrast, X-ray studies, conducted within the last few years, have focused on searching for a ~ 3.55 keV sterile neutrino signal believed to be detected first in stacked spectra of galaxy clusters (Malyshev et al., 2014; Sonbas et al., 2015). Sterile neutrinos are a dark matter candidate, a type of neutrino that does not interact with ordinary matter (including other neutrinos) but could annihilate to create photons in the keV range (Hamann & Hasenkamp, 2013). Curiously, these studies do not find any significant evidence for this 3.55 keV line, even though the signal (if it exists) should be comparatively strong (compared to known astrophysical signals) inside of dwarf galaxies.

Before the launch of the *Chandra X-ray Observatory* and *XMM-Newton* in 1999, X-ray sources were not detected in dwarf galaxies. This was for three principal reasons. Firstly, based on considerations discussed above (age, distance, central stellar density, globular cluster frequency, low galaxy mass), X-ray binaries were not expected to be detected in large numbers (if at all), so X-ray observations of dwarf galaxies were rare. For example, the relationship between the number of bright XRBs and Galaxy mass is observed to be $N_X(> 10^{37})$ erg/s = 142.9 ± 8.4 sources per 10^{11} $M_\odot$ (Gilfanov, 2004). A typical dwarf galaxy mass of $\sim 10^{6-7}$ $M_\odot$ would suggest < 1 source in these galaxies. Secondly, the best X-ray telescope of the time was *ROSAT*, which lacked the sensitivity and angular resolution of recent telescopes necessarily to effectively probe dwarf galaxy environments. Only a handful of *ROSAT* observations of dwarf galaxies were conducted, for which the results were similar to those on the Fornax Dwarf Spheroidal Galaxy: Gizis et al. (1993) detected a number of X-ray sources in its direction,

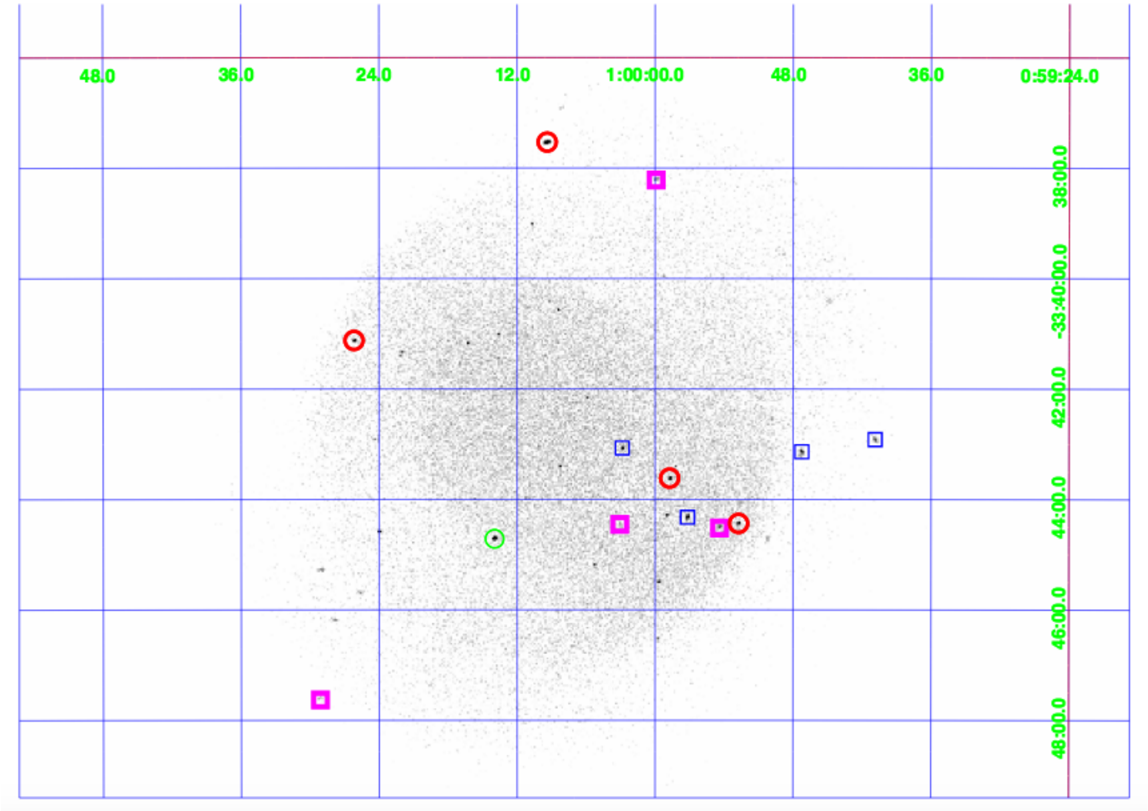

Figure 1.8: X-ray image of the Sculptor Dwarf Spheroidal galaxy. The brightest X-ray sources within the galaxy are highlighted. Thick circles (red) identify X-ray sources that were matched with evolved (post-main sequence) counterparts, thin circle (green) identifies a source that has a fainter optical counterpart, thick squares (purple) identify sources that lack high-confidence optical counterparts, while the thin squares (blue) identifies a background galaxy. This figure was adapted from Maccarone et al. (2005).

consistent with the predicted number of background active galactic nuclei (AGN). Finally, in the mid-1990s, only about a dozen Milky Way satellite dwarf galaxies were known, a number which has more than doubled since then (McConnachie, 2012).

In the *Chandra* era, studies of a number of dwarf galaxies have been conducted using more sensitive and higher angular resolution telescopes, and they we summarize them briefly here. The first detailed study was done by Maccarone et al. (2005) in the Sculptor dwarf spheroidal galaxy using *Chandra* observations. Surprisingly, as can be seen in Figure 1.8, roughly 9 sources were detected in Sculptor, of which 5 appeared to be low-mass X-ray binaries. In particular, 4 of these sources appeared to have evolved, red giant counterparts, which lends credence to the hypothesis that these sources should have evolved donors with low duty cycles. However, at the time of observation the only optical counterpart to this study involved a relatively shallow image, leaving the true classification of these LMXB candidates in doubt (Schweitzer et al., 1995). Ramsay & Wu (2006) searched for X-ray sources in the Sagittarius and Carina dSph galaxies. Sagittarius dSph has a similar star-formation history to Sculptor dSph, however it is at a low Galactic latitude with high extinction and is therefore more difficult to study in visible light. Carina dSph has more recent ($\sim$ 7 Gyr ago) star formation and less extinction, but it is further away. Both the Sagittarius and Carina observations yielded a large number of X-ray detections, however based on X-ray hardness ratios (the ratio of X-ray flux in lower energy X-ray bands to higher energy X-ray bands) the vast majority of these sources were expected to be background AGN. Between the two galaxies, about 10 total candidate LMXBs were found. However, that paper had no optical or infrared followup to better classify these candidates.

More recently, studies have observed a variety of dwarf spheroidal galaxies using *XMM-Newton*. A deep survey of the Fornax dSph was conducted by Nucita et al. (2013). As with Sagittarius and Carina, a large number of sources were found, roughly a dozen of which were classified as potential LMXBs. Unfortunately, optical counterparts were limited to those available from the 2MASS catalog. A number of these X-ray sources appear to be associated with Fornax's globular clusters (Nucita et al., 2013). Four nearby dwarf spheroidal galaxies were studied recently through archival data by Manni et al. (2015). In addition to classification using hardness ratios, they also used comparison with optical catalogues. With this analysis, they found a handful ($<$ 5 total) of sources in the selected galaxies (Draco, Leo I, Ursa Major II, and Ursa Minor dSphs). They identified a central X-ray source in the Ursa Minor dSph which correlates with a radio source, providing potential evidence for an IMBH at the galaxy's core. The Draco dSph galaxy was also studied in detail by Sonbas et al. (2015). They found, with a more recent observation, no evidence for actively accreting LMXBs but potential evidence for quiescent LMXB sources. *HST* observations were independently made of the same field, as shown in Figure 1.9. Unfortunately, however, the authors indicate that these images are too shallow and in too few filters to be of use for classifying LMXBs.

Clearly, despite nearly a half-century of satellite-based X-ray astronomy, the population of XRBs in the galaxies of the Local Group remains fertile ground for additional study. In this thesis, I have explored three galaxies of the Local Group (Sculptor Dwarf Spheroidal, Andromeda, Milky Way) using different techniques in order to understand the XRB population of each galaxy. In Sculptor Dwarf, I have followed up on the previous study of its XRB population using newer archival *Chandra* observations, as well as archival *Spitzer* photometry and Gemini photometry and spectroscopy. These combined observations allow us to deeply investigate the nature of the brightest sources in the direction of Sculptor Dwarf. For M31, I have developed a new machine learning-based classification scheme for identifying X-ray binary candidates in the galaxy using the deep *Chandra* catalogue of its bulge and sections of its disk. Using a random forest classification algorithm, we can identify new XRB candidates and remove contaminating foreground and background sources. Finally, within the Milky Way I have used results from the second release of the *Gaia* satellite to more accurately measure distances to XRBs within the Milky Way using the parallax of candidate counterparts. This enables us to use the position of XRBs within the Milky Way to understand the relationship between XRBs and the Milky Way's structure.

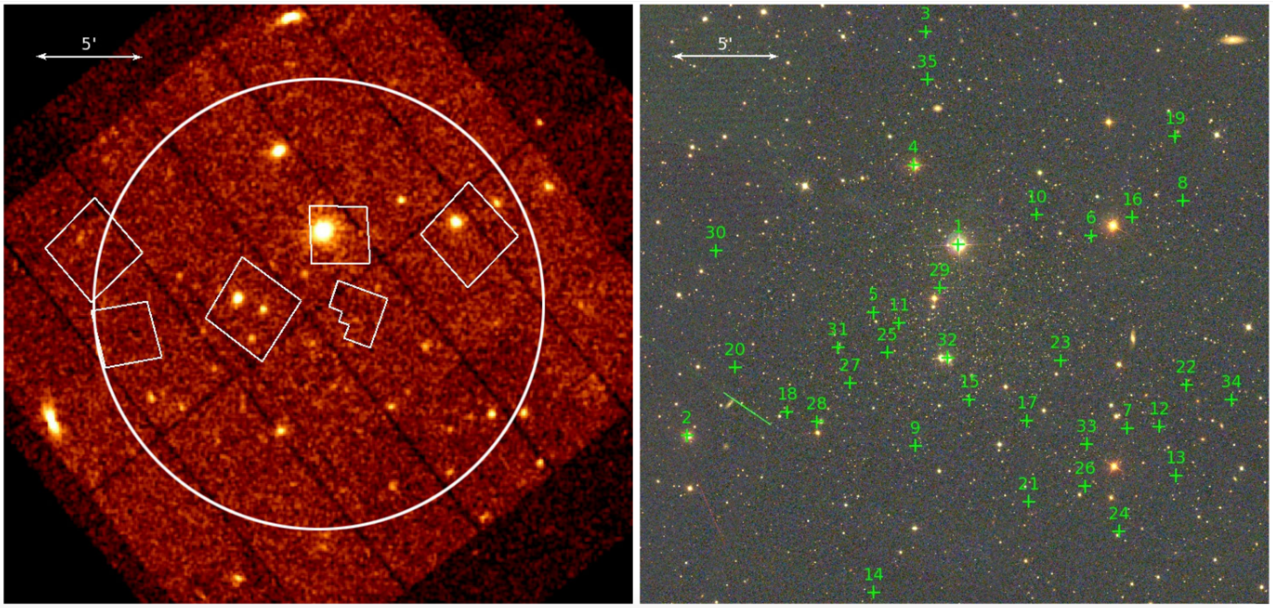

Figure 1.9: *Left:* XMM-Newton image of Draco dSph galaxy. The half-light radius of the galaxy is shown by the white circle (radius 10'), while *HST* fields are also shown in white. *Right:* Combined *u*, *g*, and *r* SDSS image of Draco, with the X-ray sources identified with green crosses. These images depict the same field. These images were; taken from Sonbas et al. (2015).

Bibliography

Aaronson, M., & Mould, J. 1980, The Astrophysical Journal, 240, 804, doi: `10.1086/158293`

Abramowski, A., Aharonian, F., Ait Benkhali, F., et al. 2014, Physical Review D, 90, 112012, doi: `10.1103/PhysRevD.90.112012`

Agar, J. R. R., & Barmby, P. 2013, AJ, 146, 135, doi: `10.1088/0004-6256/146/5/135`

Baade, W., & Zwicky, F. 1934, Proceedings of the National Academy of Science, 20, 254, doi: `10.1073/pnas.20.5.254`

Bahramian, A., Heinke, C. O., Sivakoff, G. R., & Gladstone, J. C. 2013, The Astrophysical Journal, 766, 136, doi: `10.1088/0004-637X/766/2/136`

Bahramian, A., Heinke, C. O., Sivakoff, G. R., et al. 2014, The Astrophysical Journal, 780, 127, doi: `10.1088/0004-637X/780/2/127`

Barmby, P., Huchra, J. P., Brodie, J. P., et al. 2000, AJ, 119, 727, doi: `10.1086/301213`

Bellazzini, M., Pasquali, A., Federici, L., Ferraro, F. R., & Pecci, F. F. 1995, ApJ, 439, 687, doi: `10.1086/175208`

Bergström, L. 2012, Annalen der Physik, 524, 479, doi: `10.1002/andp.201200116`

Bildsten, L., & Deloye, C. J. 2004, The Astrophysical Journal, 607, L119, doi: `10.1086/421844`

Blaauw, A. 1961, Bull. Astron. Inst. Netherlands, 15, 265

Bonnivard, V., Combet, C., Daniel, M., et al. 2015, Monthly Notices of the Royal Astronomical Society, 453, 849, doi: `10.1093/mnras/stv1601`

Bowyer, S., Byram, E. T., Chubb, T. A., & Friedman, H. 1965, Science, 147, 394, doi: `10.1126/science.147.3656.394`

Bradt, H. V., Rothschild, R. E., & Swank, J. H. 1993, A&AS, 97, 355

Bullock, J. S. 2010. `https://arxiv.org/abs/1009.4505`

Caldwell, N., Schiavon, R., Morrison, H., Rose, J. A., & Harding, P. 2011, AJ, 141, 61, doi: `10.1088/0004-6256/141/2/61`

Camenzind, M. 2007, Compact objects in astrophysics : white dwarfs, neutron stars, and black holes, doi: `10.1007/978-3-540-49912-1`

Chandrasekhar, S. 1931, ApJ, 74, 81, doi: `10.1086/143324`

Clark, G. W. 1975, ApJ, 199, L143, doi: `10.1086/181869`

Cohen, J. G. 2004, The Astronomical Journal, 127, 1545, doi: `10.1086/382104`

Dehnen, W., & King, A. 2006, Monthly Notices of the Royal Astronomical Society: Letters, 367, L29, doi: `10.1111/j.1745-3933.2005.00132.x`

Done, C., Gierliński, M., & Kubota, A. 2007, The Astronomy and Astrophysics Review, 15, 1, doi: `10.1007/s00159-007-0006-1`

Fabbiano, G. 2010, in American Institute of Physics Conference Series, ed. V. Kalogera & M. van der Sluys, Vol. 1314, 318–325

Flannery, B. P., & van den Heuvel, E. P. J. 1975, A&A, 39, 61

Fragos, T., Kalogera, V., Willems, B., et al. 2009, The Astrophysical Journal, 702, L143, doi: `10.1088/0004-637X/702/2/L143`

Frank, J., King, A., & Raine, D. J. 2002, Accretion Power in Astrophysics: Third Edition (Cambridge University Press)

Gehrels, N., Chincarini, G., Giommi, P., et al. 2004, ApJ, 611, 1005, doi: `10.1086/422091`

Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. 1962, Physical Review Letters, 9, 439, doi: `10.1103/PhysRevLett.9.439`

Gilfanov, M. 2004, Monthly Notices of the Royal Astronomical Society, 349, 146, doi: `10.1111/j.1365-2966.2004.07473.x`

Gizis, J. E., Mould, J. R., & Djorgovski, S. 1993, Publications of the Astronomical Society of the Pacific, 105, 871, doi: `10.1086/133248`

Graham, A. W., & Guzmán, R. 2003, AJ, 125, 2936, doi: `10.1086/374992`

Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2002, Astronomy and Astrophysics, 391, 923, doi: `10.1051/0004-6361:20020826`

Grindlay, J. E. 1987, in IAU Symposium, Vol. 125, The Origin and Evolution of Neutron Stars, ed. D. J. Helfand & J.-H. Huang, 173–184

Grindlay, J. E. 1988, in IAU Symposium, Vol. 126, The Harlow-Shapley Symposium on Globular Cluster Systems in Galaxies, ed. J. E. Grindlay & A. G. D. Philip, 347–363

Hamann, J., & Hasenkamp, J. 2013, J. Cosmology Astropart. Phys., 10, 44, doi: `10.1088/1475-7516/2013/10/044`

Harris, W. E. 1996, The Astronomical Journal, 112, 1487, doi: `10.1086/118116`

Harrison, F. A., Craig, W. W., Christensen, F. E., et al. 2013, ApJ, 770, 103, doi: `10.1088/0004-637X/770/2/103`

Heinke, C. O. 2010, in American Institute of Physics Conference Series, Vol. 1314, American Institute of Physics Conference Series, ed. V. Kalogera & M. van der Sluys, 135–142

Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F., & Collins, R. A. 1968, Nature, 217, 709, doi: `10.1038/217709a0`

Holberg, J. B. 2005, in Bulletin of the American Astronomical Society, Vol. 37, American Astronomical Society Meeting Abstracts, 205.01

Hubble, E. P. 1929, ApJ, 69, doi: `10.1086/143167`

Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Nature, 370, 194, doi: `10.1038/370194a0`

Ivanova, N., Justham, S., Nandez, J. L. A., & Lombardi, J. C. 2013, Science, 339, 433, doi: `10.1126/science.1225540`

Jansen, F., Lumb, D., Altieri, B., et al. 2001, A&A, 365, L1, doi: `10.1051/0004-6361:20000036`

Jarrett, T.-H., Chester, T., Cutri, R., et al. 2000, AJ, 120, 298, doi: `10.1086/301426`

Kahabka, P. 2002, arXiv e-prints, astro. `https://arxiv.org/abs/astro-ph/0212037`

Kazantzidis, S., Moore, B., & Mayer, L. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 327, Satellites and Tidal Streams, ed. F. Prada, D. Martinez Delgado, & T. J. Mahoney, 155

Kim, D., & Fabbiano, G. 2004, The Astrophysical Journal, 611, 846, doi: `10.1086/422210`

Kim, D.-W., Fabbiano, G., Brassington, N. J., et al. 2009, The Astrophysical Journal, 703, 829, doi: `10.1088/0004-637X/703/1/829`

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82, doi: `10.1086/307643`

Kormendy, J. 1985, ApJ, 295, 73, doi: `10.1086/163350`

Kraan-Korteweg, R. C., & Lahav, O. 2000, A&A Rev., 10, 211, doi: `10.1007/s001590000011`

Kroupa, P. 2001, MNRAS, 322, 231, doi: `10.1046/j.1365-8711.2001.04022.x`

Kroupa, P., Famaey, B., de Boer, K. S., et al. 2010, Astronomy & Astrophysics, 523, A32, doi: `10.1051/0004-6361/201014892`

Kundu, A., Maccarone, T. J., & Zepf, S. E. 2007, The Astrophysical Journal, 662, 525, doi: `10.1086/518021`

Lane, R. R., Kiss, L. L., Lewis, G. F., et al. 2010, MNRAS, 406, 2732, doi: `10.1111/j.1365-2966.2010.16874.x`

Lasota, J.-P. 2001, New Astronomy Reviews, 45, 449, doi: `10.1016/S1387-6473(01)00112-9`

Leigh, N., Sills, A., & Knigge, C. 2011, MNRAS, 416, 1410, doi: `10.1111/j.1365-2966.2011.19136.x`

Lewin, W. H. G., van Paradijs, J., & van den Heuvel, E. P. J. 1997, X-ray Binaries (Cambridge University Press)

Liebert, J., Young, P. A., Arnett, D., Holberg, J. B., & Williams, K. A. 2005, ApJ, 630, L69, doi: `10.1086/462419`

Livio, M. 1997, in Astronomical Society of the Pacific Conference Series, Vol. 121, IAU Colloq. 163: Accretion Phenomena and Related Outflows, ed. D. T. Wickramasinghe, G. V. Bicknell, & L. Ferrario, 845

Lyne, A. G., & Lorimer, D. R. 1994, Nature, 369, 127, doi: `10.1038/369127a0`

Maccarone, T. J., Kundu, A., Zepf, S. E., Piro, A. L., & Bildsten, L. 2005, Monthly Notices of the Royal Astronomical Society: Letters, 364, L61, doi: `10.1111/j.1745-3933.2005.00106.x`

Maccarone, T. J., Yukita, M., Hornschemeier, A., et al. 2016, MNRAS, 458, 3633, doi: `10.1093/mnras/stw530`

Malyshev, D., Neronov, A., & Eckert, D. 2014, Physical Review D, 90, 103506, doi: `10.1103/PhysRevD.90.103506`

Manni, L., Nucita, A. A., De Paolis, F., Testa, V., & Ingrosso, G. 2015. `https://arxiv.org/abs/1509.01076`

Mateo, M. L. 1998, ARA&A, 36, 435, doi: `10.1146/annurev.astro.36.1.435`

McConnachie, A. W. 2012, The Astronomical Journal, 144, 4, doi: `10.1088/0004-6256/144/1/4`

McConnachie, A. W., & Irwin, M. J. 2006, Monthly Notices of the Royal Astronomical Society, 365, 1263, doi: `10.1111/j.1365-2966.2005.09806.x`

Nucita, A., De Paolis, F., Manni, L., & Ingrosso, G. 2013, New Astronomy, 23-24, 107, doi: `10.1016/j.newast.2013.03.003`

Perryman, M. A. C., Brown, A. G. A., Lebreton, Y., et al. 1998, Astronomy and Astrophysics

Piro, A. L., & Bildsten, L. 2002, The Astrophysical Journal, 571, L103, doi: `10.1086/341341`

Podsiadlowski, P. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 401, RS Ophiuchi (2006) and the Recurrent Nova Phenomenon, ed. A. Evans, M. F. Bode, T. J. O'Brien, & M. J. Darnley, 63

Podsiadlowski, P., Pfahl, E., & Rappaport, S. 2005, Binary Radio Pulsars, 328

Podsiadlowski, P., Rappaport, S., & Pfahl, E. D. 2002, The Astrophysical Journal, 565, 1107, doi: `10.1086/324686`

Rahmani, S., Lianou, S., & Barmby, P. 2016, MNRAS, 456, 4128, doi: `10.1093/mnras/stv2951`

Ramsay, G., & Wu, K. 2006, A&A, 459, 777, doi: `10.1051/0004-6361:20065959`

Reshetnikov, V. P. 2000, Astrophysics, 43, 145, doi: `10.1007/BF02683977`

Revnivtsev, M., Postnov, K., Kuranov, A., & Ritter, H. 2011, Astronomy & Astrophysics, 526, A94, doi: `10.1051/0004-6361/201014726`

Robitaille, T. P., & Whitney, B. A. 2010, ApJ, 710, L11, doi: `10.1088/2041-8205/710/1/L11`

Sbordone, L., Bonifacio, P., Buonanno, R., et al. 2007, Astronomy and Astrophysics, 465, 815, doi: `10.1051/0004-6361:20066385`

Schweitzer, A. E., Cudworth, K. M., Majewski, S. R., & Suntzeff, N. B. 1995, The Astronomical Journal, 110, 2747, doi: `10.1086/117727`

Shapley, H. 1938, Harvard College Observatory Bulletin, 908, 1

Siegel, M. H., Dotter, A., Majewski, S. R., et al. 2007, The Astrophysical Journal, 667, L57, doi: `10.1086/522003`

Sonbas, E., Rangelov, B., Kargaltsev, O., Dhuga, K. S., & Hare, J. 2015. `https://arxiv.org/abs/1505.00216`

Sparke, L. S., & Gallagher, III, J. S. 2007, Galaxies in the Universe (Cambridge University Press)

Stiele, H., & Kong, A. K. H. 2018, MNRAS, 475, 4911, doi: `10.1093/mnras/sty106`

Stiele, H., Pietsch, W., Haberl, F., et al. 2011, A&A, 534, A55, doi: `10.1051/0004-6361/201015270`

Strader, J., Smith, G. H., Larsen, S., Brodie, J. P., & Huchra, J. P. 2009, AJ, 138, 547, doi: `10.1088/0004-6256/138/2/547`

Swartz, D. A., Ghosh, K. K., Tennant, A. F., & Wu, K. 2004, ApJS, 154, 519, doi: `10.1086/422842`

Swartz, D. A., Soria, R., & Tennant, A. F. 2008, The Astrophysical Journal, 684, 282, doi: `10.1086/587776`

Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371, doi: `10.1146/annurev-astro-082708-101650`

Tutukov, A., & Yungelson, L. 1973, Nauchnye Informatsii, 27, 70

van den Bergh, S. 1959, Publications of the David Dunlap Observatory, 2, 147

—. 2000, The Galaxies of the Local Group (Cambridge University Press)

van den Heuvel, E. P. J., & Heise, J. 1972, Nature Physical Science, 239, 67, doi: `10.1038/physci239067a0`

van Speybroeck, L., Epstein, A., Forman, W., et al. 1979, ApJ, 234, L45, doi: `10.1086/183106`

Verbunt, F., & Lewin, W. H. G. 2006, Globular cluster X-ray sources, ed. W. H. G. Lewin & M. van der Klis, 341–379

Voges, W., Aschenbach, B., Boller, T., et al. 1999, A&A, 349, 389

Vulic, N., Gallagher, S. C., & Barmby, P. 2016, MNRAS, 461, 3443, doi: `10.1093/mnras/stw1523`

Vulic, N., Hornschemeier, A. E., Wik, D. R., et al. 2018, ApJ, 864, 150, doi: `10.3847/1538-4357/aad500`

Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, in Proc. SPIE, Vol. 4012, X-Ray Optics, Instruments, and Missions III, ed. J. E. Truemper & B. Aschenbach, 2–16

Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2014, ApJ, 789, 147, doi: `10.1088/0004-637X/789/2/147`

White, N. E., & Ghosh, P. 1998, ApJ, 504, L31, doi: `10.1086/311568`

White, III, R. E., Sarazin, C. L., & Kulkarni, S. R. 2002, ApJ, 571, L23, doi: `10.1086/341122`

Willman, B. 2010, Advances in Astronomy, 2010, 1, doi: `10.1155/2010/285454`

Chapter 2

Multiwavelength survey of Sculptor Dwarf X-ray Sources

2.1 Introduction

X-ray binaries (XRBs) are rare systems consisting of a black hole or neutron star accreting material from a companion. In old stellar populations, binaries where the compact object accretes from a low-mass companion overflowing its Roche Lobe (low-mass X-ray binaries, or LMXBs) make up the brightest X-ray sources, with X-ray luminosities $L_X > 10^{34}$ erg s^{-1} (Psaltis, 2004; Heinke, 2010). At these luminosities, the X-ray spectrum is often characterized by a soft, blackbody-like thermal component that is attributed to the accretion disk, with contributions from a harder non-thermal component attributed to a hot corona of material around the compact object (Done et al., 2007). At longer wavelengths (UV through to visible), emission from the companion star is often mixed with a blue continuum of reprocessed light from the accretion disk, as well as bright emission lines due to excitation of hydrogen-like or helium-like atoms by the high temperatures of accretion flow (van Paradijs & McClintock, 1995). This implies that XRBs will tend to show a photometric excess (relative to what is expected for the companion star) in bluer UV-visible filters (Heinke, 2010). Many X-ray sources are observed to experience transient behaviour, where bright outbursts occur between long periods of quiescence. This transient behaviour is typically attributed to instability in the accretion disk due to varying viscosities from hydrogen ionization (Lasota, 2001).

2.1.1 XRB Production

X-ray binaries can be formed in stellar populations through two main pathways. In the field, XRBs are formed primordially from progenitor binary systems that survive the supernova event of the accretor. In areas of high stellar density, such as globular clusters (GCs), XRBs can be formed through dynamical encounters. For example, a compact object may be swapped into a binary system in an exchange encounter (Hills, 1991). Additionally, a compact object passing near two stars may remove enough energy from the motions of the two stars to cause them to be bound into a binary configuration (Verbunt et al., 1987; Verbunt & Lewin, 2006). If stellar density is high enough, physical collisions between stars may create new binary systems

(Fabian et al., 1975; Ivanova et al., 2006, 2008). As a result, XRBs exist in GCs at a rate roughly 100 times the Galactic field per unit mass (Clark, 1975; Verbunt & Lewin, 2006).

The efficiency of globular clusters at producing XRBs has led to the suggestion that many, or most XRBs are formed within globular clusters and then captured by their host galaxy through dynamical processes (Grindlay, 1988; White et al., 2002). However, investigating if globular clusters are supplying XRBs to their host galaxies requires detailed knowledge of the frequency and intensity of interactions between a host galaxy and its GC system. Investigations of the radial distribution of XRBs tend to suggest that globular clusters are not supplying XRBs to their host galaxies; LMXBs tend to trace a galaxy's stellar mass rather than its GC distribution (Gilfanov, 2004).

It is generally expected that XRB production should peak in a population roughly 1 Gyr after the peak of star formation (White & Ghosh, 1998). Steady-state accretion would require that a low-mass X-ray binary (LMXB) with a companion of roughly 1 $M_\odot$ should be consumed by its accretor within 100 Myr. However, this expected LMXB lifetime can be altered by a number of factors. First, many LMXBs are not persistently bright, and are observed to undergo long periods of quiescence (Piro & Bildsten, 2002). This may require that LMXBs are quiescent in excess of 75% of their accretion cycle, implying that there should be a very large number of quiescent LMXBs (qLMXBs) for each bright one. Investigations of the populations of XRBs in globular clusters suggest that this is the case—the number of persistent, bright XRBs in all Galactic globular clusters is $\sim$10, while the number of quiescent or transient XRBs found in a handful of these largest clusters (e.g., Terzan 5, M62, 47 Tuc) is at least $\sim$50 (Heinke et al., 2005, 2006; Fragos et al., 2009; Bahramian et al., 2014). Second, a number of systems are expected to evolve to a mass-transferring state much later than what is predicted by peak star formation. Ultracompact systems and systems which began as intermediate mass X-ray binaries but evolved to a lower companion mass through mass transfer are both expected to have longer lifetimes (Bildsten & Deloye, 2004; Podsiadlowski et al., 2002). Additionally, some currently existing persistently bright systems may have had no mass transfer at earlier epochs until the donor had evolved on the subgiant branch to a point that it began filling its Roche lobe. This is a potential explanation for systems that seem to exhibit no quiescence, such as Sco X-1 (Revnivtsev et al., 2011). Some systems, known as the very faint X-ray transients (VFXTs), exhibit very faint outbursts of 10^{34-36} erg s^{-1} and have low time-averaged accretion rates that are not explained by traditional outburst accretion models and would result in different accretion lifetimes (Degenaar & Wijnands, 2009, 2010; Heinke et al., 2015; Shaw et al., 2017).

The long-time evolution of individual systems is challenging to understand in both globular clusters and the field, because the age of any particular system is unknown. Measurements of system age are usually constrained through companion type and stellar evolutionary phase, neutron star magnetic field strength, system orbital period or mass transfer rate (see, e.g., Kundu et al. (2003) and references therein). In globular clusters, populations of primordial XRBs are contaminated by new XRBs created through dynamical interactions (Heinke et al., 2003; Pooley et al., 2003). In addition, globular clusters can potentially destroy wider orbit systems that might otherwise evolve to be XRBs in the present day through dynamical encounters (Verbunt & Freire, 2014). On the other hand, in the Galactic field, the age determination of individual XRB systems is complex given the presence of active star formation, interactions with members of the Local Group, high extinction through the disk, and uncertain distance estimates.

2.1.2 XRB Populations in Dwarf Galaxies

Dwarf galaxies are ideal candidates for studying the long-term population of XRBs. Some dwarf galaxies host globular clusters of their own and can be interacting strongly with their host galaxies. Sagittarius Dwarf Spheroidal (Sgr dSph) is one such example, as it is in the process of being tidally disrupted by the Milky Way (Sbordone et al., 2007). Palomar 12, a globular cluster originally thought to be associated with the Milky Way, is now believed to have been tidally stripped from the Sgr dE and is part of the Sgr tidal stream (Cohen, 2004). In principle, this process could provide a mechanism for transferring XRBs from a dwarf galaxy to its host. Additionally, M54 is generally thought to be a GC that has sunk to Sgr dSph's core through dynamical friction (Siegel et al., 2007). It is possible that this may be a mechanism for XRBs to be transferred from a GC to its dwarf host galaxy. However, most dwarf spheroidal galaxies are much further away from the Milky Way, beyond 40–50 kpc, and lack globular clusters (McConnachie, 2012).

Dwarf galaxies tend to have relatively simple star formation histories compared to other nearby galaxies, often characterized by a brief star formation event early in their histories with a small tail of more recent star formation (Monkiewicz et al., 1999). They also have central stellar densities that are roughly two orders of magnitude lower than those of globular clusters. The rate at which XRBs are formed through dynamical interactions is proportional to the square of the stellar density ρ^2 (Bahramian et al., 2013; Verbunt & Hut, 1987). This implies that dwarf galaxies will have an interaction rate approximately 4 orders of magnitude lower than typical GC values. In turn, dwarf galaxies are unlikely to have any new XRBs formed through dynamical interactions. Additionally, their lower central stellar densities means that crowding is manageable with high resolution imaging even though dwarf galaxies are on average an order of magnitude more distant than GCs.

The proximity of the many satellite dwarf galaxies to the MW means that transient LMXBs can potentially be discovered in quiescence (below 10^{33} erg s^{-1}), if observations are sufficiently deep. In addition, the presence of LMXBs can help independently constrain the size of dwarf galaxy dark matter halos, since the supernova kick given to an LMXB progenitor will generally exceed the velocity dispersion of stars in the population (Dehnen & King, 2006). Many LMXBs receive a natal kick during the formation of the compact object, which could in principle eject the system from a dwarf galaxy. However, a dark matter halo could permit a dwarf galaxy to retain XRBs that it would have otherwise lost to the Milky Way's tidal field. (Podsiadlowski et al., 2005; Dehnen & King, 2006).

Since the advent of the *Chandra* era, a number of studies have been conducted to look for LMXB candidates in dwarf galaxies. Carina, Sagittarius, Fornax, Leo I, Ursa Major II, and Ursa Minor dwarf galaxies have all been targeted with either *Chandra* or *XMM-Newton* (Ramsay & Wu, 2006; Nucita et al., 2013; Manni et al., 2015). However, these studies have generally either had low spatial-resolution or no multiwavelength counterpart matching. More recently, two separate studies of the X-ray sources in Draco dSph were conducted by Sonbas et al. (2016) and Saeedi et al. (2016) using *XMM-Newton*. These surveys both identified four LMXB candidates through matching to the Sloan Digital Sky Survey (SDSS), Wide-Field Infrared Survey Explorer (*WISE*), Two Micron All-Sky Survey (2MASS), and other optical/infrared surveys.

2.1.3 Sculptor Dwarf Spheroidal Galaxy

The Sculptor Dwarf Spheroidal (dSph) Galaxy is ideal as a low-density counterpart to globular clusters. Sculptor is one of the closest dwarf galaxies, and is at a favourable Galactic latitude (−83°) outside the plane of the Milky Way (McConnachie, 2012). It lacks globular clusters of its own, and has a relatively simple star formation history. Analysis of both *Hubble Space Telescope* (HST) and synthetic CMDs have suggested that Sculptor has a predominantly ancient stellar population and a smaller population of intermediate age stars (Monkiewicz et al., 1999; Dolphin, 2002; Tolstoy et al., 2003). A more recent analysis has shown that this intermediate age population tends to be concentrated towards Sculptor's core, and can be described with a simple, smoothly decreasing star formation rate (SFR) ending around 7 Gyr ago (de Boer et al., 2012).

In this paper we investigate the nature of bright X-ray binary candidates in Sculptor in order to constrain the size and luminosity of its XRB population. Maccarone et al. (2005) surveyed the galaxy using 21 6–ks *Chandra* pointings, and identified 9 X-ray sources with sufficient counts to accurately identify position and search for optical counterparts. These sources were matched to the optical catalogue of Schweitzer et al. (1995), and one source was ruled out as a background galaxy, while five were identified as LMXB candidates, with X-ray luminosity $> 10^{33}$ erg s^{-1}. In order to determine the nature of these candidates in detail, we combine new non-simultaneous Gemini GMOS imaging (2016) and spectroscopy (2008), along with *Spitzer* photometry (2008) and a second epoch of *Chandra* imaging. Gemini spectroscopy permits us to directly separate contaminating active galactic nuclei (AGN) and foreground stars from objects within Sculptor, while *Spitzer* and Gemini photometry allow us to look for long-wavelength counterparts to X-ray sources that are associated with Sculptor's population. We use the mid-infrared AGN selection wedge of Stern et al. (2005) to indicate whether an individual X-ray source is likely to be an AGN.

For this analysis, we assume that Sculptor Dwarf is at a distance of 86 kpc, with a heliocentric velocity $v_{\odot} = 111.4$ km s^{-1}, and a spectroscopic redshift $z = 0.000372$ (McConnachie, 2012). We use the IRSA dust maps and the Schlafly & Finkbeiner (2011) conversion to obtain a value of $A_V = 0.0484$ for the Milky Way foreground extinction in the direction of Sculptor. We further use the empirical relations of Bahramian et al. (2015) and Foight et al. (2016) to obtain a foreground column density of $N_{\rm H} = 9.0 \times 10^{19}$ cm^{-2} for Sculptor. For this analysis we assume that the Sculptor Dwarf Galaxy has negligible internal gas and dust. Single-dish Parkes observations of HI in the direction of Sculptor give it a total HI mass of $\sim 2 \times 10^4 M_{\odot}$, low amongst Local Group dwarfs for which HI measurements exist (Bouchard et al., 2003; McConnachie, 2012).

2.2 Data

2.2.1 *Chandra* Data Reduction

There are two sets of *Chandra* observations of Sculptor in the *Chandra* Data Archive. We summarize these observations in Table 2.1. For this analysis, we considered the count rates from the first epoch, first studied by Maccarone et al. (2005). Additionally, we considered in detail a new observation (ObsID: 9555) for both count rates and spectral analysis. The new observation was made with the Advanced CCD Imaging Spectrometer (ACIS-S) in faint mode with an effective frame time of 3.1 s. We reduced the observation using CALDB version 4.7.2, the August 2008 time-dependent gain file, and CIAO 4.8 (Fruscione et al., 2006). We reprocessed the observation using the `chandra_repro` script and obtained a new level 2 events file. Exposure-corrected images were created using `fluximage` in the default "broad" band (0.5 - 7.0 keV). Sources were detected on these images using `WAVDETECT`.

Spectral extraction was performed on all sources with > 100 source counts using `ACISEXTRACT` (version: 2016sep22, MARX: 5.3, CIAO 4.9, CALDB 4.7.3, IDL 8.5) (Broos et al., 2010). We then fit each spectrum using XSPEC 12.9.0 (Arnaud, 1996). Sources were grouped to have 25 counts per bin where possible and fit using χ^2 statistics. For each source, we compute a model-dependent flux hardness ratio calculated as $F_{0.5-2.0\ \mathrm{keV}}/F_{2.0-8.0\ \mathrm{keV}}$. The model used is XSPEC's power law model `pegpwrlw` multiplied by `tbabs` photoelectric absorption. We used default abundances, $N_{\mathrm{H}} = 9.0 \times 10^{19}$ cm^{-2}, and permit the photon index and normalization of the power law to vary. Additionally, we calculate a model-independent count hardness ratio using the count rate C obtained from the `show rate` command in XSPEC, calculated as $C_{0.5-2.0\ \mathrm{keV}}/C_{2.0-8.0\ \mathrm{keV}}$. Sources previously detected by Maccarone et al. (2005) that lacked sufficient statistics were fit using `cstat` instead (Cash, 1979; Arnaud, 1996). For purposes of comparison between epochs, we present the background-subtracted count rate for each source in each epoch using the values from Wang et al. (2016). SD X-6 and X-10 were not observed by Wang et al. (2016), so we computed their count rates using aperture photometry with the CIAO tool `dmstat`. For X-6, we used a circular extraction region of 3″and a background annulus with inner radius 3″and outer radius 8″. For X-10, the source is significantly extended, so we used a circular extraction region of 16″and a background annulus with inner radius 16″and outer radius 40″. The positions, previous identifications, and X-ray properties are summarized in Tables 2.2 and 2.3. If the source was detected in our observations, we report the position derived from `WAVDETECT`. Otherwise, we report the position of the source from Maccarone et al. (2005). Matching these positions to catalogues at other wavelengths is discussed below, in Section 2.2.5.

2.2.2 Gemini Imaging Data Reduction

We obtained observations of the core of Sculptor using Gemini GMOS-S in the r and Hα bands (PID: GS-2016B-Q66, Dates: 2016-08-14 & 2016-08-15, PI: R. Arnason). The central 5.5×5.5 arcmin of Sculptor was imaged using 20×500 s Hα exposures and 8×180 s r exposures. The observations were split between two nights, only one of which had sufficiently good seeing to be usable. Therefore, our final images are comprised of 10×500 s in Hα and 4×180 s in r. From use of the IRAF task `imexamine`, we estimate the FWHM of these images to be roughly 4 pixels, combined with 0.16″per pixel gives a FWHM of 0.64″. The reduction was performed using PyRAF v2.2dev. We derived normalized flatfields and bias frames using GIFLAT and GBIAS. All images were reduced using GIREDUCE, and we corrected a defect on the GMOS-S detector using an observatory-supplied script (G. Gimeno, private communication). The images were then mosaiced using GMOSAIC and coadded using IMCOADD to create the final images, which had 2×2 binning and a 0.16″ pixel size.

To detect and extract sources, we used SExtractor 2.19.5. Source extraction was done with the following settings: 5″ fixed aperture, a 3σ detection threshold, and a minimum of

Table 2.1: Summary of *Chandra* X-ray observations of Sculptor Dwarf Spheroidal Galaxy. All observations were made using ACIS-S.

ObsID	Date	Exposure Time ks	PI	Mode
4698	2004-04-26 08:33:34	6.06	Maccarone	VFAINT
4699	2004-05-07 04:39:34	6.27	Maccarone	VFAINT
4700	2004-05-17 10:26:45	6.1	Maccarone	VFAINT
4701	2004-05-30 22:16:43	6.07	Maccarone	VFAINT
4702	2004-06-12 14:19:32	5.88	Maccarone	VFAINT
4703	2004-06-27 20:08:18	5.88	Maccarone	VFAINT
4704	2004-07-12 01:08:02	5.91	Maccarone	VFAINT
4705	2004-07-24 07:51:23	5.83	Maccarone	VFAINT
4706	2004-08-04 11:08:39	6.08	Maccarone	VFAINT
4707	2004-08-17 04:52:59	5.88	Maccarone	VFAINT
4708	2004-08-31 05:46:02	5.88	Maccarone	VFAINT
4709	2004-09-16 03:47:43	6.09	Maccarone	VFAINT
4710	2004-10-01 10:50:52	5.88	Maccarone	VFAINT
4711	2004-10-11 14:32:47	5.88	Maccarone	VFAINT
4712	2004-10-24 01:32:41	6.08	Maccarone	VFAINT
4713	2004-11-05 02:28:24	6.07	Maccarone	VFAINT
4714	2004-11-20 18:48:32	6.04	Maccarone	VFAINT
4715	2004-12-05 06:40:06	5.68	Maccarone	VFAINT
4716	2004-12-19 08:46:48	6.01	Maccarone	VFAINT
4717	2004-12-29 23:47:24	6.07	Maccarone	VFAINT
4718	2005-01-10 04:08:34	6.06	Maccarone	VFAINT
9555	2008-09-12 00:11:10	49.53	Zepf	FAINT

8 pixels above the threshold required for source detection. We derived an r-band magnitude calibration through comparison with r-band catalogues in the Dark Energy Survey (DES). We cross-matched sources in our imaging with those in the DES images, and derived a constant magnitude correction that we apply to match SExtractor's magnitudes to those measured by DES. For the Hα data, we fit a line through the main sequence of the r vs. Hα - r colour-magnitude diagram. We then calculated a constant magnitude correction such that the main sequence had H$\alpha - r = 0$. As such, our Hα magnitudes are relative and not an absolute calibration. We calculate the photometric uncertainties using the SExtractor defaults. The resulting catalogue was matched to the X-ray catalogue in TOPCAT using a $1''$ tolerance.

2.2.3 Gemini Spectroscopic Data

We have obtained observations of two fields in Sculptor with GMOS-S in Multi-object spectroscopy (MOS) mode (PID: GS-2008B-Q-25, Date: 2008-09-21 PI: S. Zepf). The spectra were taken using the B600 G5323 grating, with 3 exposures of 420 s per field. The spectra have a resolution of approximately 8 Å (as measured from the FWHM of the arc lamp lines) and a wavelength range of roughly 4000–7000 Å. Four of the sources discovered by Maccarone et al. (2005) were observed in this program. Each of the three exposures were taken was slightly different central wavelengths in order to remove the GMOS chip gaps from the final spectra. We derived normalized spectral flats and bias frames with GSFLAT and GBIAS. Arc, standard, and science images were created using GSREDUCE. Each exposure was filtered for cosmic rays using GEMCRSPEC. The images were wavelength calibrated using GSWAVELENGTH, and sky subtracted using GSSKYSUB. We extracted 1D spectra from the 2D spectra using GSEXTRACT, and then stacked the exposures together with GEMSCOMBINE. The stacked exposures were then flux calibrated using GSSTANDARD and the spectrum of the standard star (LTT1020). The resulting final spectra were analyzed and fit using SHERPA 4.9.0.

2.2.4 *Spitzer* Data

Mid-infrared images of Sculptor were obtained in *Spitzer* Cycle 5 (PID: 50314, PI: P. Barmby). Infrared Array Camera (IRAC; Fazio et al., 2004) observations were made in 2008 December and covered the galaxy at wavelengths of 3.6, 4.5, 5.8, and 8.0 μm in a 0.4×0.7 deg map with 5×12 s dithered observations per sky position. The expected 5σ point source detection limits are roughly 10 μJy at 3.6 and 4.5 μm and 100 μJy at 5.8 and 8.0 μm. The IRAC point-spread function FWHM is approximate 1.8–2.0$''$. Multiband Imager and Photometer for *Spitzer* (MIPS; Rieke et al., 2004) observations were made in 2008 August in scan-map mode, using medium scan speed with half-array offsets to cover an area of 0.4×1.2 deg. While data were collected in all three of the MIPS bands, here we focus on only the 24 μm data as the spatial resolution of the 70 and 160 μm bands is low compared to the other wavelengths of interest. The expected 5σ point source detection limit is roughly 800 μJy at 24 μm, with a PSF FWHM of $6''$.

A detailed description of data processing and full catalogues are in a forthcoming paper (Barmby et al., 2018, in preparation); here we give a brief summary. IRAC images were processed using pipeline S17.0.1. The Level 1 basic calibrated data (BCD) files were cleaned with custom cleaning scripts to remove artifacts. Images were mosaiced using IRACPROC post-BCD Processing package 4.0, removing transient events and fixed-pattern background noise (Schuster et al., 2006). Source extraction of IRAC mosaics was performed using SExtractor and the following settings: detection threshold of 1.5σ, 5 pixel minimum area, apertures of $2.46''$, $3.66''$, and $6.08''$ with aperture corrections from the IRAC Instrument Handbook (Center, 2015). Aperture photometry was done with IRAF/apphot, and photometric uncertainties were estimated using apphot's standard formula. MIPS images were processed using the MIPS Data Reduction tool. Point sources were extracted for all MIPS bands using the PSF-fitting program StarFinder with model PSFs from Engelbracht et al. (2007). MIPS sources were matched to IRAC 3.6 μm sources using TOPCAT. Most ($\sim$ 80%) MIPS sources within the IRAC coverage area had IRAC counterparts (those without were most often located near bright stars), but as expected for normal stars, most IRAC sources do not have MIPS counterparts.

2.2.5 Matching

In order to classify sources detected in our X-ray catalogue, we have carefully considered cross-matching between the *Chandra* detections, our own catalogues, and external catalogues retrieved from ViZieR. We have also considered cross-matching with photometry from the Dark Energy Survey, which provides coverage of Sculptor in *griz*Y filters (Abbott et al., 2018). *Chandra*'s absolute positional uncertainty is $\sim 0.8''$. For ordinary on-axis point sources, we expect that associations should match within this tolerance. However, we chose a larger tolerance of $2''$ to account for several factors within the dataset. Firstly, we lacked an absolute astrometry correction. Secondly, a number of the sources in the field lie off-axis, resulting in greater positional uncertainty. Thirdly, at least one of the X-ray sources appears to be extended (see the section on SD X-10, below), and its optical counterpart is extended over several arcseconds. In addition, a number of the X-ray sources from Maccarone et al. (2005) are relatively faint or are detected only as upper limits, and as such, the positional accuracy is reduced. We expect that point sources in our catalogue should match within $\sim 0.8''$, and permit a higher tolerance only for these extended, faint, or off-axis sources. As delivered, the Gemini GMOS r and Hα images were misaligned in World Coordinate System (WCS) space. Therefore, the astrometry was aligned by matching bright sources in the individual filter images to bright sources in the *Spitzer* 3.6 μm images with TOPCAT (Taylor, 2005) and deriving an average shift. We selected the 100 brightest sources in each image to derive the shift. This correction only accounts for misalignments due to translation, and not rotation or scale corrections; we used the standard deviation of the difference between the *Spitzer* and r and Hα images to estimate the size of the positional uncertainty due to rotation or scale. For both the r and Hα images, we find that this uncertainty is roughly $\sim 0.2''$.

2.3 Analysis

We re-detect six of the nine sources from Maccarone et al. (2005) in our new *Chandra*observation, and we identify four new sources that have > 100 source counts, making them suitable for spectral extraction. Where possible, we use the position obtained in our new observation with `WAVDETECT`. For the sources not detected in the new observation, we use the position from

Maccarone et al. (2005). Note that we find the position of one source, SD X-8, to be different from its reported location in Maccarone et al. (2005) (see Section 2.3.8, below). To classify sources, we use X-ray and optical spectroscopy, IR and optical photometery, and comparison with catalogues. For sources with *Spitzer* photometry, we use the AGN selection wedge of Stern et al. (2005), to identify sources as AGN. Amongst the X-ray sources in the Sculptor dwarf field, we find that two are background galaxies, seven are background AGN, and the nature of three is uncertain. In this section we discuss each of the X–ray sources individually, with the population properties explored in the following section. A summary of literature IDs, classifications from the literature, and the classification derived in this work for each source is given in Table 2.2. Each source's X-ray properties are summarized in Table 2.3, and the properties of its optical/mid-IR counterparts, if any, are given in Table 2.4.

2.3.1 SD X-1

The X-ray source at this position has existing counterparts in the literature. Previous quasi-stellar object (QSO) catalogues have identified both a radio (Tinney et al., 1997) and optical counterpart (Perlman et al., 1998). Previous optical spectroscopy by Perlman et al. (1998) identified a single broad optical feature at roughly 5250 Å and a velocity width of roughly 3800 km s^{-1}. The relative isolation of the line led to its identification as Mg II. This line, plus the observed radio flux, led to the subsequent identification of this source as a BL Lac with redshift $z = 0.875$.

The GMOS-S spectrum of this object, shown in Figure 2.1, is more complicated. The doublet structure of the 5250 Å feature is more clearly revealed. Additionally, two narrow absorption features centred at 6566 Å and 4863 Å also appear in this spectrum. These lines have observed wavelengths consistent with Hα and Hβ at the systemic velocity of Sculptor Dwarf. The line widths are marginally larger than the spectral resolution, though there is overlap with the measured width of the CuAr lamp spectra. If these narrow absorption lines are associated with the same object as the BL Lac, then their presence is difficult to explain, as their rest-frame wavelengths are not associated with any known absorption features in AGN. In order to understand the nature of this object, we have examined available catalogues with counterparts at other wavelengths.

Aside from the initial radio and optical counterparts in QSO catalogues, there are substantial archival observations of the source, though none are simultaneous. A spectral energy distribution (SED) of available archival data is plotted in Figure 2.2. Radio observations of the source made with the Australia Telescope Compact Array (ATCA, 2011 July/August) and the NRAO VLA Sky Survey (NVSS, 1996 Sept–Oct) seem to suggest time-variability in the radio flux measured at 1.4 and 2.0 GHz, respectively. GALEX images do not show any counterpart at X-1's location. *Spitzer* photometry reveals a very red object (see Figure 2.5) that is very bright in redder IRAC channels (5.8 and 8.0 μm) relative to other sources in the field, but imaging shows it is possibly contaminated by a nearby object at the longest wavelengths. This source lies outside of the Stern AGN selection wedge, as shown in Figure 2.5. Ground based optical photometry gives the source $BVRI \sim 20$.

The most plausible explanation is that the observed source is a blend between the BL Lac object (contributing the broad optical emission lines, radio, IR, and X-ray emission) and a foreground object in Sculptor contributing the narrow absorption lines. However, this explanation is clearly incomplete. It does not provide a reason why the absorption lines were not observed in the previous ground-based spectrum, or their specific origin. We also lack a plausible candidate for the source of the Balmer absorption lines if in Sculptor: the age of Sculptor's stellar populations would prohibit normal strong Balmer line sources, such as A stars.

One possibility for the origin of these lines is that they are absorption caused by a cloud of H I gas in Sculptor along the line of sight to the AGN. We can estimate the required column density using the curve of growth and assuming the cloud is optically thin. Using the measured equivalent widths of the lines, we then calculate the column density as:

$$N = 1.13 \times 10^{20} \frac{EW}{\lambda^2 f} \text{ cm}^{-2} \tag{2.1}$$

where EW is the equivalent width in Å, λ is the rest wavelength of the line in Å, and f is the line oscillator strength (Spitzer et al., 1974; Frisch, 1972). Using $f_{H\alpha} = 0.64$, $f_{H\beta} = 0.12$, and measuring $EW_{H\alpha} = 2.31 \pm 0.3$ Å and $EW_{H\beta} = 2.02 \pm 0.2$ Å from the lines, we derive column densities $N_{H\alpha} = 9.4 \times 10^{12}$ cm^{-2} and $N_{H\beta} = 8.1 \times 10^{13}$ cm^{-2}. These column densities represent the density of hydrogen atoms in the $n = 2$ state along the line of sight. We can compare this to the neutral hydrogen density (presumed to be ground state hydrogen) observed by Bouchard et al. (2003). The column density at the location of SD X-1 using Parkes to detect the 1.4 GHz spin-flip transition falls on a contour with $N_{n=1} = 1 \times 10^{18}$. The ratio of the $N_{n=2}$ column density to the $N_{n=1}$ column density implies, assuming the Boltzmann distribution, temperatures of ~20,000 K. It is difficult to explain such a high ISM temperature inside Sculptor, especially given the dearth of potential ionizing sources in the galaxy that could heat the ISM. However, since the ratio compares a spin-transition to an ordinary atomic electron transition, it is likely that the assumption of the Boltzmann distribution is inaccurate. Ideally, the required column density of neutral hydrogen should be calculated using a measurement from the Lyman transitions. As such, based on existing data, we cannot conclude whether these lines are due to intervening H I gas in Sculptor.

Another possibility for the origin of the Balmer lines is that they are caused by a star with high proper motion that lies in the slit of the GMOS spectrum but is outside of it in the Perlman et al. (1998) spectrum. The GMOS pre-imaging of the source shows a single point source at the resolution of the observation (roughly 0.5″). To look for a counterpart, we inspected archival USNO catalogues, shown with GMOS pre-imaging and *Spitzer* imaging in Figure 2.3. The USNO images do not contain any plausible source that could be a high proper-motion object appearing in the slit of the GMOS spectrum. We note, as a caveat, that these images are shallower and have poorer angular resolution than the GMOS imaging. Based on the GMOS spectrum, examination of the SED, and comparison with catalogues, we conclude that SD X-1 is most likely a background AGN with some foreground object in Sculptor causing foreground absorption, though the source of this absorption is unknown.

2.3.2 SD X-2

This X-ray source also has an existing counterpart in the literature. Tinney et al. (1997) identify the source as QJO100-3341 with an optical spectrum showing lines at ~4500 Å, ~7900 Å, and 8000 Å. They identify these lines as Mg II, Hβ and [O III], with a calculated redshift of 0.602.

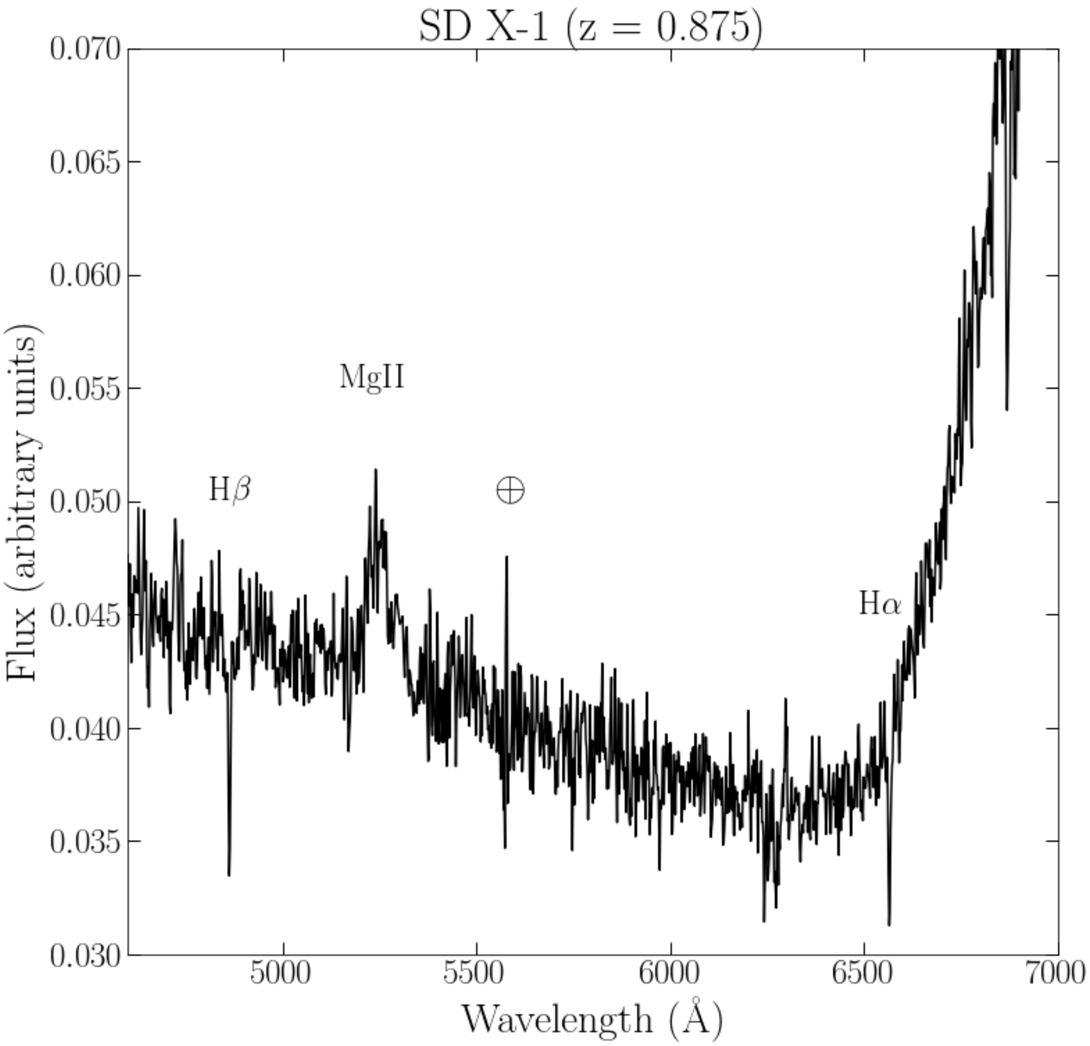

Figure 2.1: GMOS-S MOS observed-frame spectrum of SD X-1. Note the presence of a broad feature at ~ 5400 Å, which we associate with the same feature identified as Mg II in the spectrum of Perlman et al. (1998). Additionally, note the two narrower features, not found in the Perlman et al. (1998), at approximately the rest wavelengths of Hα and Hβ. The rise in the continuum redward of 6500 Å is instrumental, and the feature at ~ 5600 Å is telluric.

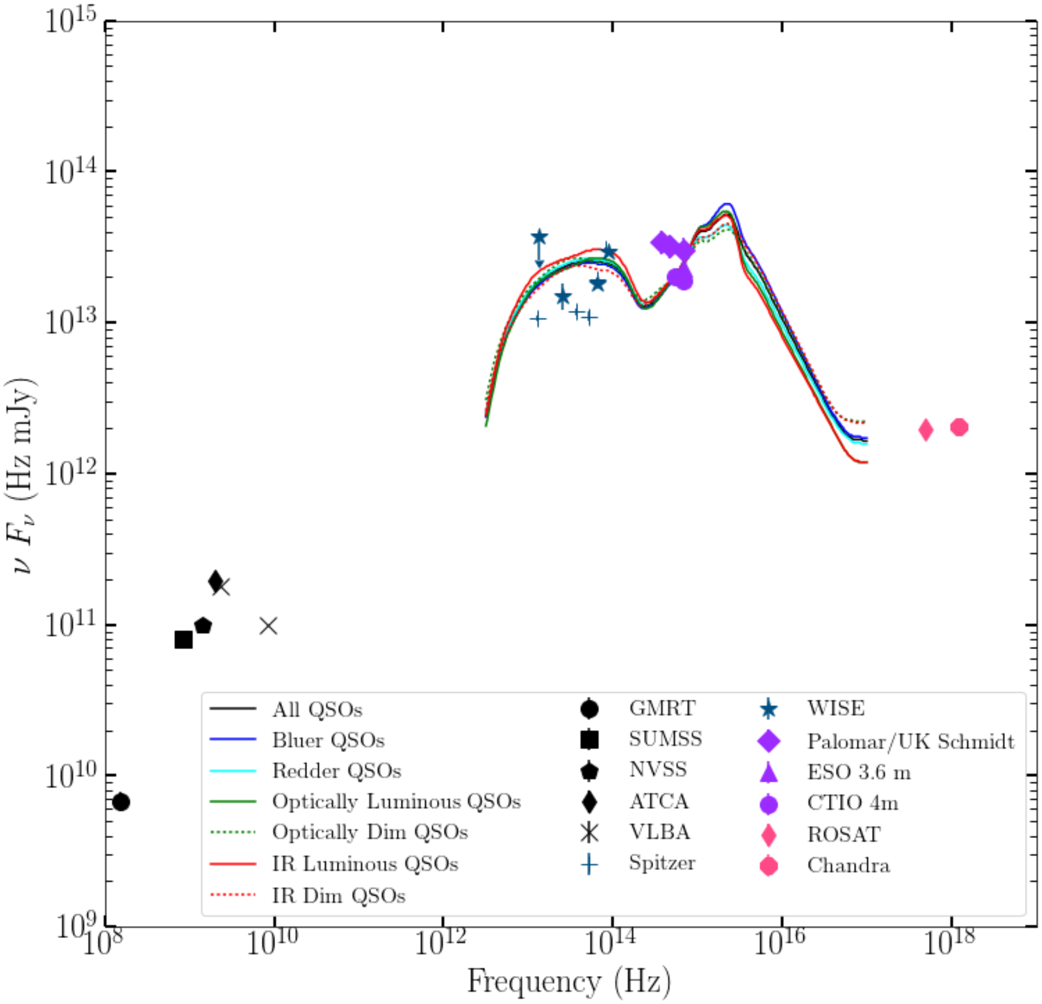

Figure 2.2: Non-simultaneous spectral energy distribution (SED) for SD X-1 and all counterparts. Error bars are included from catalogues where estimated, and upper limits are indicated when available. Curves plotted over the SED are composite QSO spectra from Richards et al. (2006). Curves are normalized to have the same V-band flux density as the value measured by CTIO 4m V band photometry (Schweitzer et al., 1995).

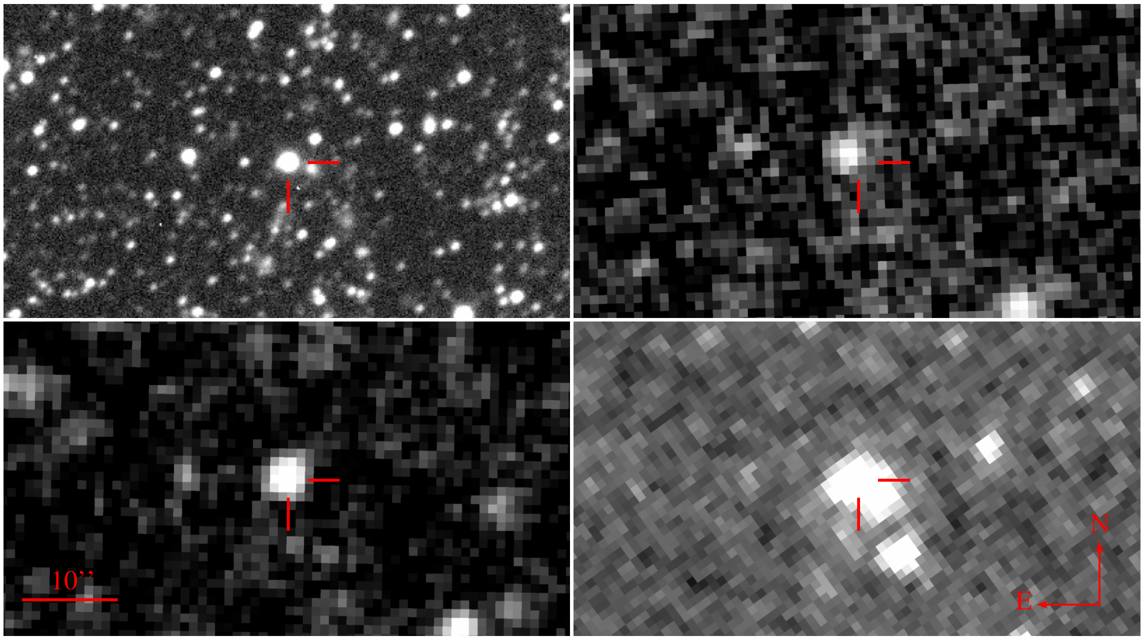

Figure 2.3: Finding charts for SD X-1. Clockwise from top left: Gemini GMOS-S r-band pre-imaging (2008), USNO AAOR image (1997), USNO SRCJ image (1982), and *Spitzer* IRAC 3.6 μm image. Up is North and left is East. The X-ray position of SD X-1 is bracketed by two lines. To the resolution of the GMOS pre-imaging, SD X-1's r-band counterpart appears to be a point-source. Note the lack of any obvious high proper motion interlopers in the USNO imaging, while the *Spitzer* image shows a nearby source which may be contaminating photometry at IR wavelengths.

In our optical spectrum, shown in Figure 2.4, we detect an emission line at 4485 Å, as well as new lines at 5476 and 5975 Å. We compare the observed-frame positions of these lines to strong lines in the SDSS template spectra. The positions of these new lines allow us to self-consistently identify the line at 4485 Å as C IV, making the other lines C III] and Fe II. From these identifications, we derive a new redshift of $z = 1.895$. Based on the broad lines in the optical spectrum, we classify this source as an AGN.

2.3.3 SD X-3

This X-ray source has no existing counterpart in the literature, aside from detections in photometric surveys. We find that this source has an optical counterpart in our GMOS imaging with $r = 21.2$ and $\mathrm{H}\alpha = 21.1$. Additionally, in our GMOS-MOS observation we have discovered a new spectroscopic counterpart. The GMOS spectrum, shown in Figure 2.4 contains two broad lines. One line is at $\sim$ 4300Å, and the other is at at $\sim$ 5300Å. Comparing with strong lines in SDSS template spectra, we can self-consistently assign these lines to be C IV and C III]. With this tentative identification, X-3 has a derived redshift of $z = 1.795$. X-3's IR emission falls at the edge of the Stern AGN selection wedge. It also has a relatively soft X-ray spectrum. The presence of broad lines in the optical spectrum leads us to classify this source as an AGN.

2.3.4 SD X-4

SD X-4 has an existing optical counterpart in the literature. Tinney (1999) obtained an optical spectrum showing a broad absorption line quasar (BAL QSO) (ID: QJ0059-3344) with C III] and Mg II lines at ~5900 and ~8200 Å, respectively, with a redshift of $z = 2.022$. An object at the same location was found by Jones et al. (2009), but with a lower S/N spectrum and a number of weak lines identified with a redshift of $z = 0.07630$. The GMOS spectrum of this object, shown in Figure 2.4, is of considerably higher quality than either existing spectrum in the literature in the region 4000–6000 Å. We identify the same C III] line and BAL QSO features as Tinney (1999), confirming the redshift of $z = 2.022$. In addition, we further identify C IV as the absorbed feature at ~4660 Å, and find additional features at 4220 and 4800 Å, which we identify as S IV,O IV] and He II, respectively. This object's IR counterpart falls deep in the AGN selection wedge, as shown in Figure 2.5. We also identify a Gemini counterpart with $r = 19.4$ and $\mathrm{H}\alpha = 19.2$. Based on the optical spectrum and its IR colours, we classify this object as an AGN.

2.3.5 SD X-5

The optical counterpart of SD X-5 was previously identified by Maccarone et al. (2005), however it was not targeted for GMOS spectra as it has magnitude $V = 23.68$. With the Australia Telescope Compact Array (ATCA), Regis et al. (2015) identified a radio source $\sim 2''$away that could potentially be associated. X-5 also has a relatively soft X-ray spectrum. *Spitzer* matching identifies a counterpart which has colours in the Stern AGN region. The L_X/L_{opt} ratio of this object is rather high compared to other AGNs in this field, and as such this object could potentially be a qLMXB or a dusty AGN. Based on the IR colours, we tentatively classify this object as an AGN.

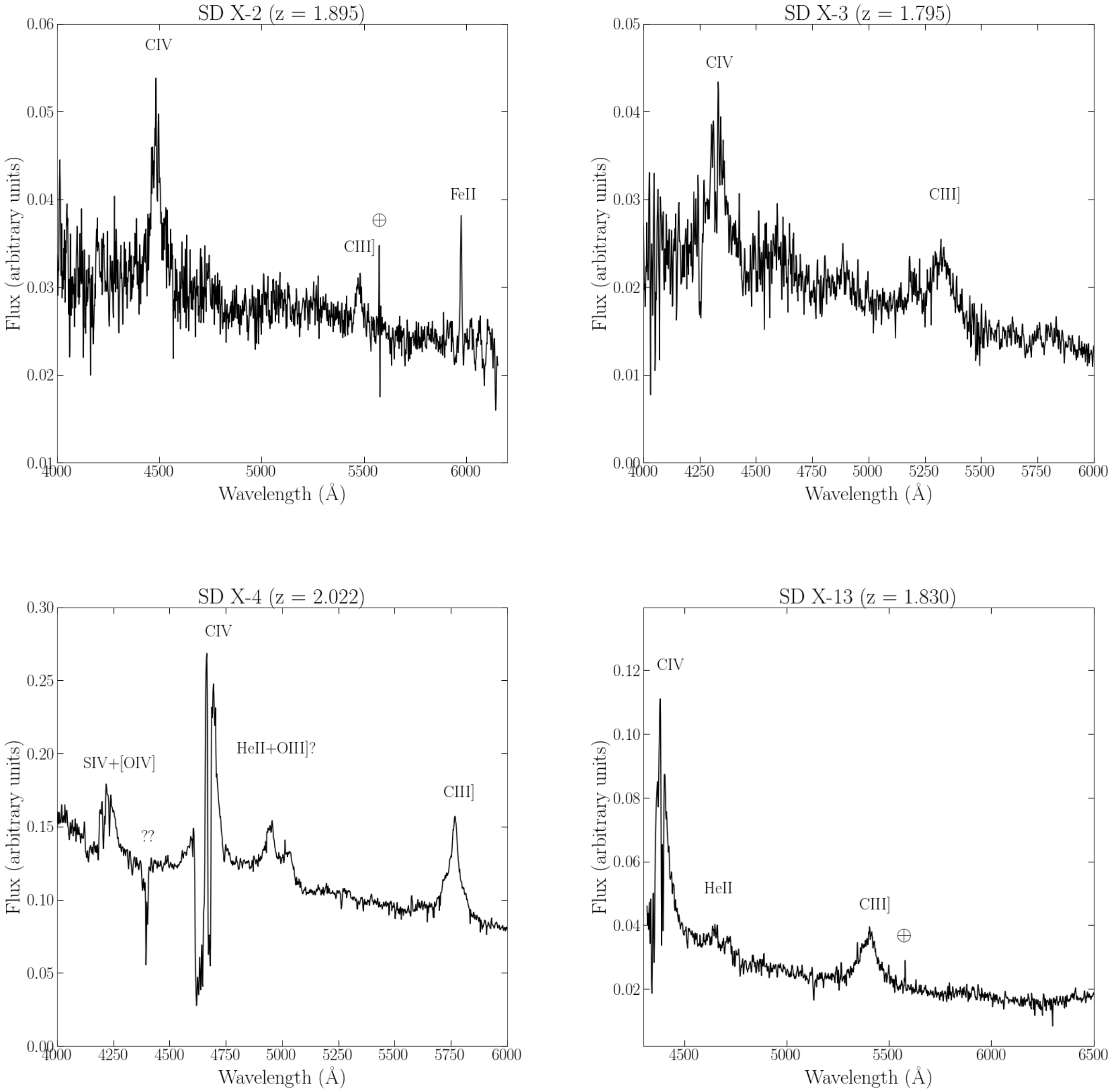

Figure 2.4: GMOS-S MOS observed-frame spectra for SD X-2 (top left), SD X-3 (top right), SD X-4 (bottom left), and SD X-13 (bottom right). The feature at ~5600 Å in the spectra of X-2 and X-13 is telluric.

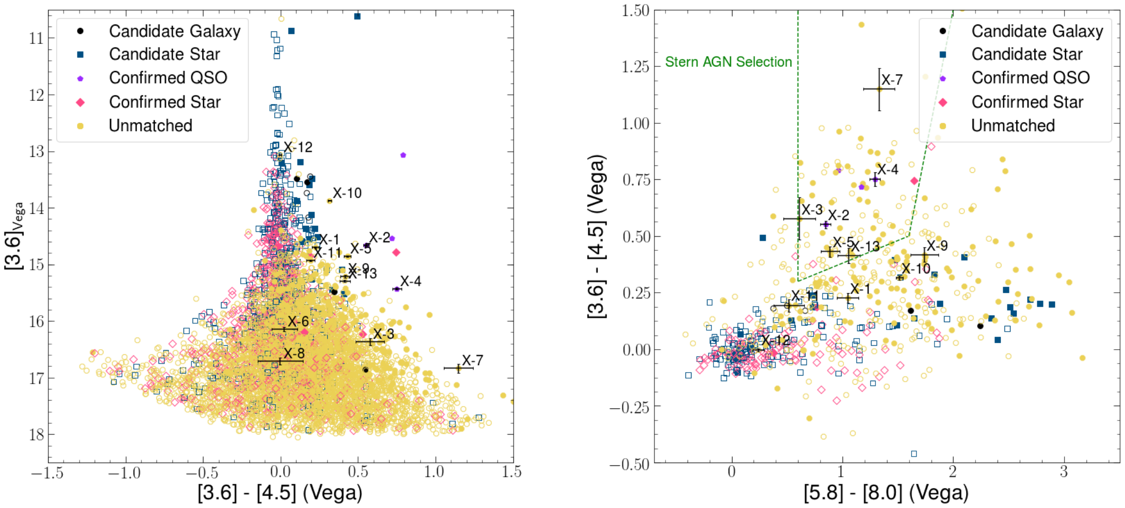

Figure 2.5: Left: *Spitzer* [3.6] μm vs [3.6] μm - [4.5] μm colour-magnitude diagram for Sculptor. Right: *Spitzer* [3.6] μm - [4.5] μm vs [5.8] μm - [8.0] μm colour-colour diagram for Sculptor. We only plot sources with a [3.6] μm magnitude uncertainty < 0.2. Sources with a MIPS 24.0 μm counterpart detected are filled, those without MIPS detections are unfilled. Matches to X-ray sources are plotted with error bars. We also plot the Stern et al. (2005) AGN selection wedge. Note that candidate and confirmed galaxies, stars and QSOs are based on matching to NED and SIMBAD.

2.3.6 SD X-6

There is no existing literature identification for SD X-6. This source is also not detected in the 2nd epoch of *Chandra* observations. It is detected in both Spitzer and Gemini GMOS imaging, though it lacks 5.8 μm and 8.0 μm detections. Its position on the Spitzer CMD puts it closer to the main sequence than AGN or background galaxy counterparts. In addition, its position on the r-Hα CMD, shown in Figure 2.6, suggests that it has less Hα-band emission relative to ordinary sources in the galaxy. It is difficult to make a definitive identification, given that this source has no followup X-ray detection, partial IRAC detection, and no spectrum. However, given the Hα dearth relative to the main sequence, it appears it is unlikely to be an XRB in Sculptor. Possible classifications could be an AGN, an AGN blended with a foreground star in Sculptor, or a foreground MW star.

2.3.7 SD X-7

SD X-7 has a match in the quasar catalogue published by Flesch (2015). The catalogue entry lists R (20.31) and B (20.37) magnitudes. However, the catalogue entry indicates that the B and R magnitudes were measured years apart. Since QSOs can show variability, the colour is unlikely to be reliable. The catalogue lists a probability of being a QSO $P_{QSO} = 0.98$, though it is unclear how the probability is derived from the optical photometry, and there is no spectroscopic classification of this object in the catalogue. It was not targeted for a GMOS spectrum. It lies outside of the field of view in the 2nd epoch of *Chandra* observations, and was therefore not detected. This source has a photometric counterpart detected by *Spitzer* but is outside the field of view of our Gemini imaging. SD X-7 has a counterpart in DES photometry with a separation of 0.075″, as shown in Figure 2.7. In the bluer CMD, X-7 appears consistent with a horizontal branch star. However, in the redder CMD, X-7 is slightly redder than the main sequence overall, suggestive of an AGN. The IR counterpart is relatively faint and red, as can be seen in Figure 2.5. Additionally, its position on the *Spitzer* colour-colour diagram puts it deep in the Stern AGN selection. Based on its IR colours, we classify X-7 as a background AGN.

2.3.8 SD X-8

This source appears to be at a slightly different location from where it was reported in Maccarone et al. (2005). `WAVDETECT` finds that, in both epochs of observations, the coordinates of this source are $\sim 2''$ away from the first reported X-ray position. We also note that this source is faint and it appears extended, possibly due to off-axis effects. This source has a relatively hard X-ray spectrum. The 3σ position found by `WAVDETECT` overlaps with a counterpart in *Spitzer* imaging. The position on the *Spitzer* CMD suggests that it is an ordinary star, though it is not detected at 5.8 μm or 24.0 μm. This mid-infrared counterpart corresponds with an optical counterpart located $\sim 2''$ away from the reported X-ray position, found by Kirby et al. (2013) and Kirby et al. (2015). We also find a Gemini GMOS counterpart to the *Spitzer* object in r and $H\alpha$, which has $r -$ Hα colour consistent with the main sequence (no excess). The Kirby et al. (2013) counterpart is a red giant with $T_{\rm eff} = 4904$ K and $\log(g) = 2.10$. If this red giant were the Roche lobe-filling donor in an XRB with a 0.5 - 10 $M_\odot$ accretor, the system would have a period of ∼weeks. This red giant was only observed once in the study of Kirby et al. (2015), so we cannot look for radial velocity variation due to an unseen companion. If the red giant is a companion in an XRB system, it is possible that the spectrum could show emission lines associated with XRBs, such as Hα. However, the spectrum of this red giant from Kirby et al. (2015), kindly provided to us by the authors, is not unusual. In particular, it does not show Hα emission, which might be expected of such a source (E. Kirby, private communication). The lack of Hα excess in either spectral or photometric measurements suggests that this red giant is unaffiliated with the X-ray source. The optical spectrum of this source also shows no evidence of AGN emission, suggesting that it is not blended with an AGN. The Dark Energy Survey catalogue identifies no counterpart except for a source consistent with location of the red giant. Based on this, we conclude that the *Spitzer* Gemini, and DES counterparts are the red giant from Kirby et al. (2013), and are not associated with the X-ray source. As such, we detect no optical/IR counterpart to X-8. The Dark Energy Survey has a limiting magnitude of $\sim$ 24 in *r*-band (Abbott et al., 2018), which should detect sources down to an absolute magnitude of $\sim$ 4.3, assuming a distance modulus of 19.67 (McConnachie, 2012). Therefore, based on the lack of an associated optical counterpart, we tentatively classify X-8 as a background AGN that is nearby on the sky to an unaffiliated red giant.

2.3.9 SD X-9

This source, previously classified as a background galaxy by Schweitzer et al. (1995), lies out of the field of the 2nd epoch of *Chandra* observations and was not detected. It also lies out of the field for GMOS imaging. It possesses an IR counterpart, with relatively red IR colours. Based on prior identifications, we maintain the classification as a background galaxy.

2.3.10 SD X-10

SD X-10 is the brightest X-ray source in the second epoch of *Chandra* observation. This source lies outside of the field covered by the S3 chip in the first epoch of observations, however it is detected on other chips and is included in the X-ray source catalogue of Liu (2011). It has no obvious literature counterparts, however its X-ray identification is complicated by `WAVDETECT` finding 2 sources in this location, spread out over a few arcseconds. This could be a true extended object, an artifact of being off-axis, or multiple sources at the same location. The source position lies out of the field of the Gemini observations, but has a *Spitzer* counterpart 2.5$''$ away from the X-ray location. This *Spitzer* counterpart has a very red IR colour placing it outside the Stern AGN selection. In the DSS images of these coordinates, we find a possibly extended source, spread over a few arcseconds, most likely a background galaxy. Based on DSS and *Chandra* both showing a bright extended object, we conclude that this source is a background galaxy.

2.3.11 SD X-11

SD X-11 has no known counterparts in the literature. It has a relatively hard X-ray spectrum. This source is also detected in the first epoch of *Chandra* data, but it was not reported by Maccarone et al. (2005), most likely due to a lack of an optical counterpart in the Schweitzer et al. (1995) catalog. At Sculptor's distance, this object would have an X-ray luminosity of $\sim 3.5 \times 10^{34}$ erg s^{-1}, plausible for an LMXB. It is detected in *Spitzer* but not Gemini imaging. As such, it was not targeted for GMOS spectroscopy. Its *Spitzer* colours lie outside the AGN selection wedge, however its colours are not consistent with those of an normal star. Since the object is detected in X-ray and IR but has no optical counterpart, it could be a dusty AGN or a high redshift galaxy. Based on the available photometry and X-ray properties, we identify this source as a candidate AGN/background galaxy.

2.3.12 SD X-12

SD X-12 has no known counterparts in the literature. It is detected in the *Spitzer* and Gemini imaging, however it is saturated in Gemini. We also identify a very bright counterpart in DES with a separation of 0.0075″. The IR counterpart has $[3.6]_{\mathrm{Vega}} = 13.07$, making it one of the brightest objects in the population. This object was detected in the first epoch of *Chandra* data, but was not identified by Maccarone et al. (2005), most likely because it was not in the Schweitzer et al. (1995) catalogue due to saturation. Its position in the colour-colour diagram (see Figure 2.5) is in a portion of the CCD occupied primarily by confirmed or candidate stars in Sculptor. However, the photometry of this counterpart revealed by DES is clearly distinct from the ordinary stellar population of Sculptor. The object is several magnitudes brighter than giant branch stars in Sculptor, as seen in the bluer CMD (see Figure 2.7 left panel). However, in redder filters, X-12's counterpart appears significantly redder than the ordinary population (see Figure 2.7 right panel). The overall brightness of this object and its large colour disparity from Sculptor's population suggests that it is likely to be a foreground star.

This object is unlikely to be an XRB or a CV, as the required optical contribution of the accretion disk creates an implausible optical/X-ray flux ratio given an implied X-ray luminosity of $\sim 3.5{\times}10^{34}$ erg s^{-1} if it is in Sculptor. Additionally, this object is observed to show significant variability - the i magnitude in the DES survey is ~ 2 magnitudes fainter than the I magnitude measured in the USNO-B1 survey. This object also has observed flaring in the ASASSN survey (Shappee et al., 2014; Kochanek et al., 2017). Based on the DES magnitudes and the presence of variability, we conclude that this source is likely to be a foreground active binary (AB) or flaring star.

2.3.13 SD X-13

SD X-13 was found by `WAVDETECT` in the second epoch of *Chandra* observations. This object was located well off-axis on the S3 chip in the first epoch of observations, and so was not reported by Maccarone et al. (2005), as its position could not be easily determined. Its *Chandra* spectrum is very soft, and it has IR colours within the Stern AGN wedge. This object is also detected in our Gemini imaging, and is shown in Figure 2.6 to have an Hα excess relative to the main sequence. Additionally, the GMOS spectrum, shown in Figure 2.4, shows three broad emission features: a strong, broad line that overlaps with the instrumental cutoff at ~4400 Å, a broad, faint double centered around 4600 Å, and a strong, broad line at 5407 Å. Through comparison with template spectra and typical strong lines, we conclude that the features at 4400, 4600, and 5407 Å are C IV, He II, and C III, respectively. From this, we derive a redshift of $z = 1.830$ and conclude that this source is a background AGN.

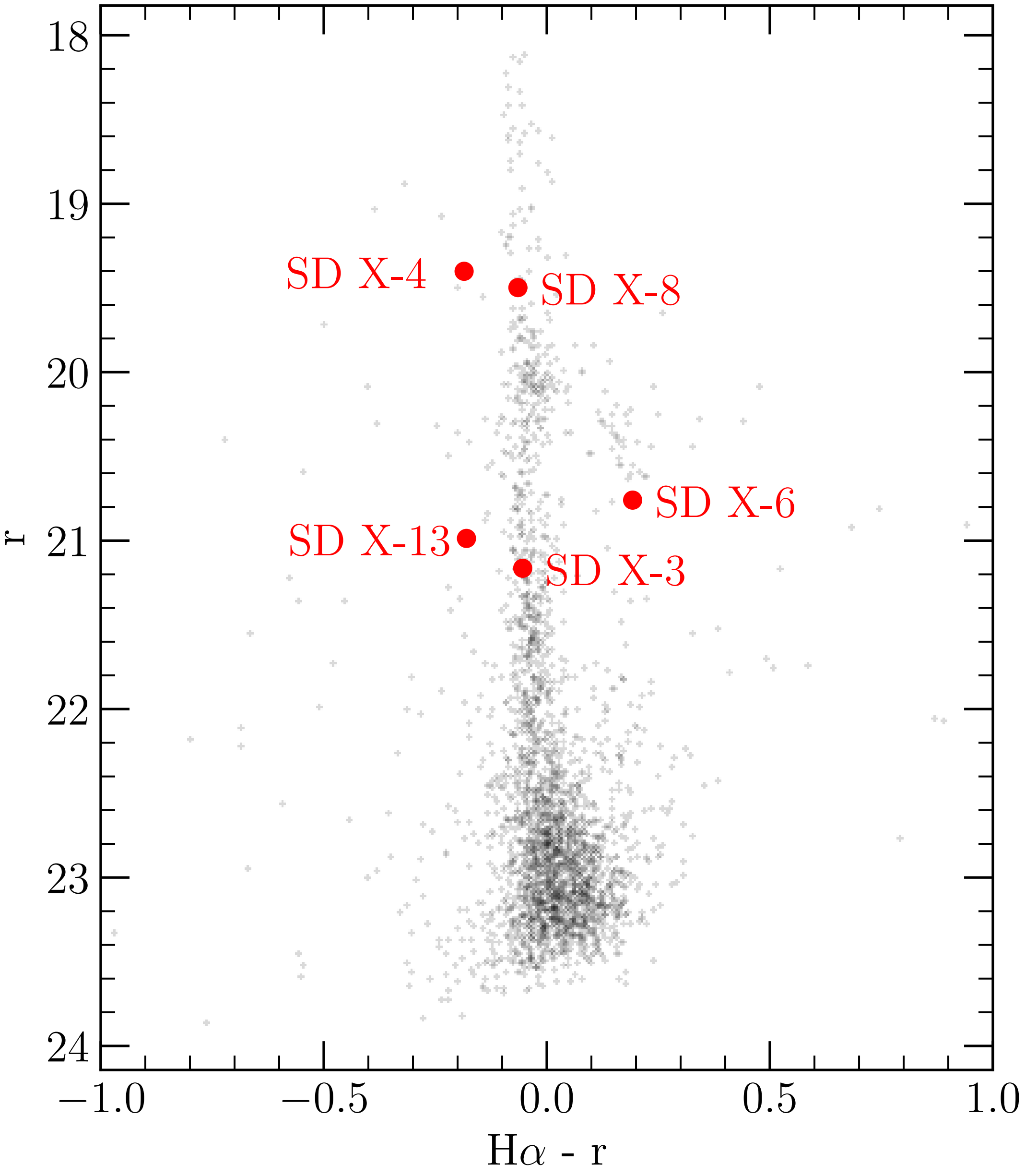

Figure 2.6: Gemini GMOS r-Hα CMD for the core (5.5 arcmin^2) of Sculptor. X-ray sources with Gemini imaging counterparts are labelled. Accreting sources are typically expected to show Hα excess relative to the main sequence, which would be on the left side of the CMD.

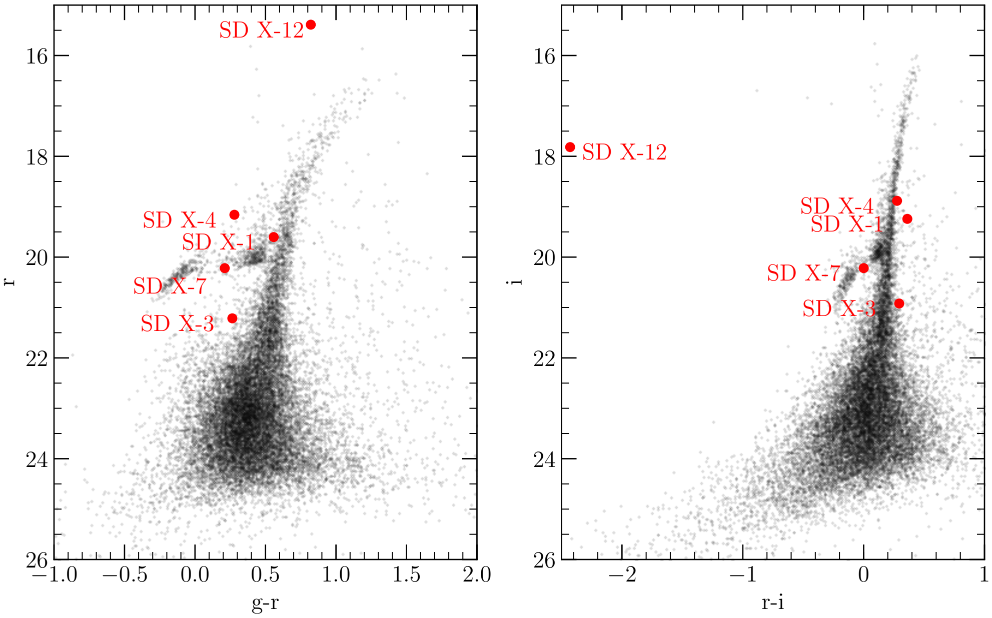

Figure 2.7: Dark Energy Survey (DES) colour-magnitude diagrams (CMDs) for the Sculptor Dwarf Galaxy. Note that all counterparts to X-ray sources except X-7 do not appear to have a consistent position relative to Sculptor's stellar population, and X-12 does not look like a Sculptor member at all.

Table 2.2: Summary of X-ray positions and literature identifications for the X-ray sources in the direction of Sculptor.

Object	RA_X	DEC_X	Other ID	Lit. Classification	Classification (TW)
SD X-1	01:00:09.39	-33:37:31.900	PKS 0057-338[1];WGAJ0100.1-3337[2]	BL Lac	BL Lac + foreground
SD X-2	01:00:26.19	-33:41:07.500	QJO100-3341[1][3]	AGN	AGN
SD X-3	00:59:58.68	-33:43:37.100	···	···	QSO
SD X-4	00:59:52.75	-33:44:26.100	QJ0059-3344[4]	QSO	BAL QSO
SD X-5	01:00:13.84	-33:44:43.300	J010014.0-334442.7[5]	···	AGN Candidate
SD X-6	01:00:03.03	-33:44:26.900	···	···	AGN/AGN+FG/FG(MW)
SD X-7	00:59:59.89	-33:38:12.200	J005959.90-333811.4[6]	QSO	AGN
SD X-8	00:59:54.38	-33:44:30.460	Scl 1008920[7][8][9];J005954.2-334429[10]	Red Giant	Unassoc Red Giant + AGN
SD X-9	01:00:29.22	-33:47:37.100	···	Background Galaxy[11]	Background Galaxy
SD X-10	01:00:27.33	-33:51:10.345	···	···	Background Galaxy
SD X-11	00:59:57.18	-33:44:19.093	···	···	AGN/Bkgd Glx Cand
SD X-12	01:00:02.85	-33:43:04.166	···	···	Flaring foreground star/AB
SD X-13	00:59:47.26	-33:43:07.106	···	···	AGN

For sources re-detected or newly detected in the 2nd *Chandra* epoch, we use the location determined from `WAVDETECT` in that observation. Otherwise, we use the position reported in Maccarone et al. (2005). [1]: Tinney et al. (1997),[2]: Perlman et al. (1998),[3]: Véron-Cetty & Véron (2010),[4]: Tinney (1999),[5]: Regis et al. (2015),[6]: Flesch (2015),[7]: Kirby et al. (2015),[8]: SIMBAD,[9]: Walker et al. (2009),[10]: Flesch (2016),[11]: Schweitzer et al. (1995)

Table 2.3: Summary of X-ray properties for sources in the Sculptor Dwarf field

Object	F_X 10^{-14} erg s^{-1}	Flux HR_X Model Dep.	Count HR_X Model Ind.	$\frac{\text{Fit Statistic}}{d.o.f.}$	d.o.f.	Epoch 1 Count Rate 10^{-3} cts/s	Epoch 2 Count Rate 10^{-3} cts/s
SD X-1	$4.7^{+0.9}_{-0.7}$	1.1 ± 0.6	4.2 ± 0.9	0.194	4	11.8 ± 0.4	4.8 ± 0.3
SD X-2*	$2.0^{+0.5}_{-0.4}$	0.8 ± 0.4	3.0 ± 0.8	0.015	62	1.9 ± 0.1	2.2 ± 0.2
SD X-3*	$1.4^{+0.3}_{-0.3}$	1.4 ± 0.8	6 ± 2	0.012	60	2.2 ± 0.1	2.4 ± 0.2
SD X-4*	$1.8^{+0.6}_{-0.4}$	0.4 ± 0.2	1.9 ± 0.5	0.019	53	1.1 ± 0.1	1.5 ± 0.2
SD X-5	$5.4^{+0.8}_{-0.7}$	0.6 ± 0.1	4.2 ± 0.7	0.098	9	9.5 ± 0.3	6.9 ± 0.4
SD X-6*	···	···	···	···	···	0.11 ± 0.08†	0.11‡
SD X-7*	···	···	···	···	···	1.7 ± 0.1	··· ± ···
SD X-8*	$0.5^{+0.4}_{-0.2}$	0.2 ± 0.3	1.1 ± 0.6	0.073	14	0.7 ± 0.1	0.5 ± 0.1
SD X-9*	···	···	···	···	···	4.9 ± 0.3	··· ± ···
SD X-10	19^{+2}_{-1}	0.9 ± 0.2	3.4 ± 0.3	0.052	23	37 ± 1	16.1 ± 0.6†
SD X-11	4^{+2}_{-1}	0.3 ± 0.2	1.0 ± 0.3	1.305	1	3.2 ± 0.2	2.5 ± 0.2
SD X-12	$4.6^{+0.6}_{-0.5}$	2.0 ± 0.7	···	0.207	10	3.9 ± 0.2	8.9 ± 0.4
SD X-13	$4.0^{+0.7}_{-0.7}$	0.7 ± 0.3	7 ± 2	0.027	6	3.2 ± 0.2	6.4 ± 0.4

Fluxes and count rates are calculated for the 0.5 - 8.0 keV band. Model dependent flux hardness ratio is calculated as $F_{0.5-2.0\ \mathrm{keV}}/F_{2.0-8.0\ \mathrm{keV}}$. Model independent count hardness ratio $C_{0.5-2.0\ \mathrm{keV}}/C_{2.0-8.0\ \mathrm{keV}}$. Sources marked with $*$ had less than 100 counts and were fitted using `cstat`. Source marked with $\star$ were either out of the field of view of the observation or had too few counts for ACISextract to obtain a spectrum. All other sources were fit using χ^2 statistics. Count rates were taken from Wang et al. (2016). Count rates marked with † were computed manually using aperture photometry in CIAO. ‡ indicates a 3σ upper limit for a non-detection.

Table 2.4: IR/visible properties of SD X-ray sources.

Object	$[3.6]_{\mathrm{Vega}}$	$[4.5]_{\mathrm{Vega}}$	$[5.8]_{\mathrm{Vega}}$	$[8.0]_{\mathrm{Vega}}$	$[24.0]_{\mathrm{Vega}}$	r_{Vega}	$\mathrm{H}\alpha_{\mathrm{Vega}}$	X-ray/IR Offset $''$
SD X-1	14.70 ± 0.01	14.47 ± 0.02	14.33 ± 0.08	13.29 ± 0.06	9.82 ± 0.05	···	···	1.04
SD X-2	14.67 ± 0.01	14.12 ± 0.01	13.34 ± 0.04	12.50 ± 0.03	8.63 ± 0.02	···	···	0.45
SD X-3	16.37 ± 0.06	15.79 ± 0.07	14.5 ± 0.1	13.9 ± 0.1	9.84 ± 0.05	21.164 ± 0.006	21.11 ± 0.02	0.16
SD X-4	15.43 ± 0.02	14.68 ± 0.02	13.66 ± 0.04	12.37 ± 0.03	8.82 ± 0.02	19.401 ± 0.001	19.216 ± 0.004	0.16
SD X-5	14.85 ± 0.01	14.42 ± 0.02	14.07 ± 0.06	13.18 ± 0.05	10.12 ± 0.06	···	···	1.54
SD X-6	16.14 ± 0.04	16.12 ± 0.07	···	···	10.38 ± 0.07	20.760 ± 0.004	20.95 ± 0.02	1.7
SD X-7	16.84 ± 0.08	15.69 ± 0.05	14.7 ± 0.1	13.42 ± 0.07	10.51 ± 0.07	···	···	0.12
SD X-8	16.71 ± 0.06	16.7 ± 0.1	···	15.4 ± 0.4	···	19.498 ± 0.002	19.433 ± 0.005	2.14
SD X-9	15.21 ± 0.02	14.80 ± 0.03	14.7 ± 0.1	12.95 ± 0.04	9.52 ± 0.04	···	···	1.05
SD X-10	13.874 ± 0.006	13.558 ± 0.008	13.20 ± 0.03	11.69 ± 0.01	7.79 ± 0.02	···	···	2.52
SD X-11	14.93 ± 0.01	14.74 ± 0.02	14.39 ± 0.09	13.9 ± 0.1	10.42 ± 0.08	···	···	0.05
SD X-12	13.068 ± 0.004	13.071 ± 0.006	13.18 ± 0.03	12.94 ± 0.04	···	···	···	0.21
SD X-13	15.29 ± 0.02	14.88 ± 0.03	14.35 ± 0.08	13.29 ± 0.06	···	20.987 ± 0.005	20.81 ± 0.01	0.84

X-ray/IR Offset was determined by matching the *Chandra* coordinates to the *Spitzer* catalogue using TOPCAT. Note that $H\alpha$ is calibrated relative such that the main sequence has $\mathrm{H}\alpha - r = 0$, rather than having an absolute calibration.

2.4 Discussion

In this study, we have completed the deepest, most thorough survey of an isolated, old stellar population with a low stellar encounter rate, and we have found that amongst the bright X-ray sources in Sculptor's direction, 7 are AGNs, 2 are background galaxies, 1 is a foreground star, and three have uncertain identifications. In the conservative case, Sculptor appears to lack any bright XRBs at the present day. Dehnen & King (2006) used the previous result of Maccarone et al. (2005) to argue that Sculptor may need a dark matter halo of 10^9 $M_\odot$ in order to retain LMXBs, unless there exists a class of LMXBs with preferentially lower natal kicks. Although low natal kick sources have been observed, these do not represent a large fraction of the LMXB population (Podsiadlowski et al., 2005). However, the absence of bright LMXBs in Sculptor would tend to imply that such large dark matter halos are unnecessary. If Sculptor is representative of dwarf galaxies in terms of present day XRB population, then there are also possible implications for interactions between a dwarf galaxy and its host or a dwarf galaxy and its globular cluster system. First, one immediate implication is that dwarf galaxies do not contaminate their host galaxies with significant quantities of XRBs when they interact with them. Additionally, if globular clusters can contaminate their host galaxy with XRBs, then Sculptor's lack of present day XRBs may suggest that it has also lacked globular clusters in the past.

2.4.1 Primordial binary contributions to observed populations

Dwarf galaxies like Sculptor can be shown to have very low stellar encounter rates. Sculptor's radial brightness profile can be reasonably described using a King model (Irwin & Hatzidimitriou, 1995). As such, we estimate the stellar encounter rate Γ using the following relation (Verbunt & Hut, 1987):

$$\Gamma = \frac{\rho_c^2 r_c^3}{\sigma} \tag{2.2}$$

where ρ_c is the central luminosity density, r_c is the core radius, and σ is the central velocity dispersion. We use $r_c = 145$ pc, based on a distance of 86 kpc and an apparent core radius size of $5.8'$ (Irwin & Hatzidimitriou, 1995; McConnachie, 2012). Additionally, we use $\sigma = 9.2$ km s^{-1} (Burkert, 2015). We use central surface brightness $\mu_c = 10.6 L_\odot/\mathrm{pc}^2$, which corresponds to $\rho_c = 0.041 L_\odot/\mathrm{pc}^3$ using the relation between projected and volume luminosities given by Djorgovski (1993). Under these assumptions, and using the normalization that 47 Tuc has $\Gamma = 1000$, Sculptor dSph has $\Gamma = 0.009$. Based on this analysis, we expect that any XRBs inside of Sculptor should be formed primordially rather than through stellar encounters. Additionally, Sculptor's relative isolation and lack of globular clusters suggests that it is unlikely to have captured XRBs through galaxy-scale interactions with the MW or GCs. Comparing with the calculations of Γ to those of Bahramian et al. (2013) (Table 4), we find that Sculptor dSph has a very small stellar encounter rate compared to Galactic GCs, most similar to that of the low-density clusters Arp 2 and Palomar 4. In general, these clusters tend to be distant, sparse, and low-mass compared to the overall population of galactic GCs, making them difficult to observe. With a stellar mass of $2.3 \times 10^6 M_\odot$, Sculptor outweighs many globular clusters by $\sim$ 1-2 orders of magnitude, and its total mass including dark matter is closer to $3 \times 10^7 M_\odot$ (Battaglia et al., 2008; McConnachie, 2012; Kimmig et al., 2015).

From our study, we see that Sculptor is essentially devoid of XRBs in the present day, which implies that it has no bright primordial binaries that have survived to the present epoch. There are a few explanations for why this may be the case. The first is that, despite being embedded in large DM haloes, natal kicks to XRB systems may be ejecting them from dwarf galaxies. Dehnen & King (2006) investigated the size of dark matter halo needed to retain XRBs based on the initial reported discovery of LMXBs inside Sculptor, with the assumption that the dark matter within the visible galaxy is 5×10^7 $M_\odot$. We can estimate the central escape velocity from Sculptor using this assumption and the following relation (Dehnen & King, 2006):

$$v_{esc}^2(r) = v_0^2 ln\frac{r_t^2 + r_c^2}{r^2 + r_c^2} \tag{2.3}$$

where r_c is the core radius, r_t is the tidal radius, and v_0 is the asymptotic circular speed defined by:

$$v_0 = 12 \ \ \mathrm{kms}^{-1} \sqrt{1 + (\frac{r_c}{1.5\mathrm{kpc}})} \tag{2.4}$$

We use $r_t = 15$ kpc and $r_c = 101$ pc to arrive at $v_{esc,0} \sim 38$ km s^{-1} (Irwin & Hatzidimitriou, 1995; Dehnen & King, 2006). Although this value is relatively low, it is comparable to the central escape velocities of a number of Galactic globular clusters (see, for example, McLaughlin & van der Marel 2005).

This order of magnitude estimate suggests that Sculptor could retain primordial binaries in its core in the same manner as a globular cluster. Therefore, the absence of bright XRBs in Sculptor would imply that the bright XRBs observed in GCs are all dynamically formed closer to the present epoch rather than primordial. The second explanation for the lack of primordial bright binaries in Sculptor is that they were never created in the first place: Sculptor's low stellar mass means that not enough primordial XRBs were created to result in some surviving to the present day. This in turn would imply that the minimum required mass for an old, isolated population to still have bright XRBs in the present day should be larger than a few 10^6 $M_\odot$. From simple mass considerations, we wouldn't expect a significant population of bright XRBs in Sculptor. Gilfanov (2004) gives an empirical relationship between the number of bright XRBs and galaxy stellar mass, with $N_X(> 10^{37})$ erg s^{-1} $= 142.9 \pm 8.4$ sources per 10^{11} $M_\odot$. Sculptor's mass of $\sim 10^6$ $M_\odot$ would scale to <1 bright XRB expected. As such, we can expect that Galactic globular clusters with an encounter rate similar to Sculptor should also be devoid of binaries, as they have much lower mass and a negligible stellar encounter rate.

2.4.2 Implications for Dwarf Galaxies

If Sculptor is representative of other dwarf galaxies in the Local Group this would imply that a large fraction of the X-ray sources in the field of nearby dwarfs are in fact unrelated to the galaxies themselves. Many of these sources are likely to be either background AGN or foreground active/flaring stars. Our search for X-ray sources suitable for X-ray spectroscopy places constraints on the X-ray luminosity function above a few 10^{34} erg s^{-1}, and in the conservative case Sculptor appears to lack any sources brighter than this limit. Two recent surveys of Draco dSph with *XMM-Newton* each identified a handful of XRB candidates. Sonbas et al. (2016) and Saeedi et al. (2016) each identify four candidate XRBs. Three of these appear to be faint sources, possibly CVs, symbiotic stars, or qLMXBs. One source has a reported $L_X = 8 \times 10^{34}$ erg s^{-1}, brighter than the limit we have investigated for Sculptor. Nucita et al. (2013) investigated Fornax dSph with *XMM-Newton*, finding in general that the number of X-ray sources was consistent with the predicted number of background sources for the area surveyed. However, they also identified two sources potentially associated with globular clusters bound to Fornax. This result appears to be consistent with Sculptor, which lacks globular clusters of its own. Manni et al. (2015) investigated four dSphs (Draco, Ursa Major II, Ursa Minor, and Leo I), also finding that the number of sources detected in the direction of each galaxy was consistent with background predictions, but noting that they could not rule out the possibility of a limited number of these sources being associated with the galaxy itself. It is interesting to note that a number of compact object-related phenomena are found preferentially in dwarf galaxies. The only repeating fast radio burst, FRB 121102, has been localized to a dwarf galaxy (Tendulkar et al., 2017) and both superluminous supernovae and long-duration gamma bursts seem to be preferentially located in dwarf galaxies (Fruchter et al., 2006; Perley et al., 2016). It is therefore curious that dwarf galaxies in the local universe, like Sculptor, do not appear to have many compact objects. FRB host galaxies do share a number of differences from DGs like Sculptor, the most prominent being that the host galaxy of FRB 121102 has a relatively high star formation rate of 0.4 $M_\odot$ yr^{-1}. This suggests that properties like active star formation may be crucial for hosting exotic compact object phenomena or, more generally, unknown phenomena with proposed compact object origins.

2.4.3 Future Studies

Dwarf galaxies like Sculptor present a unique challenge for XRB searches because of their large angular size on the sky, which means that the expected number of AGNs contaminating any dwarf galaxy is expected to be large. As such, the X-ray source population of dwarf galaxies can only be accurately characterized using deep, high-resolution multiwavelength observations, like the *Chandra*, *Spitzer*, and Gemini observations used in this study.

Although we have characterized the bright X-ray sources in this galaxy, limits on the optical sensitivity of the study have prevented us from accurately characterizing the CV population. A thorough study of Sculptor would require depths similar to those used in globular cluster surveys, so that the population of CVs can be identified and studied in detail. For example, using the CV populations of M80 and NGC 6397 as analogues to Sculptor implies that we would require an imaging depth of $R \sim 24.5$ to capture 50% of the CV population in Sculptor (Pietrukowicz, 2009; Cohn et al., 2010). In particular, CVs are often separable from the ordinary stellar population using UV or deep Hα photometry. Our Hα limits are not deep enough to capture the typical CV population, and Sculptor lacks UV observations with *HST*. A future study of Sculptor's X-ray population would seek to characterize not only the bright population of X-ray sources, but also the fainter X-ray sources contained within Sculptor's population.

Bibliography

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ArXiv e-prints. `https://arxiv.org/abs/1801.03181`

Arnaud, K. A. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes, 17

Bahramian, A., Heinke, C. O., Degenaar, N., et al. 2015, MNRAS, 452, 3475, doi: `10.1093/mnras/stv1585`

Bahramian, A., Heinke, C. O., Sivakoff, G. R., & Gladstone, J. C. 2013, ApJ, 766, 136, doi: `10.1088/0004-637X/766/2/136`

Bahramian, A., Heinke, C. O., Sivakoff, G. R., et al. 2014, ApJ, 780, 127, doi: `10.1088/0004-637X/780/2/127`

Battaglia, G., Helmi, A., Tolstoy, E., et al. 2008, ApJ, 681, L13, doi: `10.1086/590179`

Bildsten, L., & Deloye, C. J. 2004, ApJ, 607, L119, doi: `10.1086/421844`

Bouchard, A., Carignan, C., & Mashchenko, S. 2003, AJ, 126, 1295, doi: `10.1086/377312`

Broos, P. S., Townsley, L. K., Feigelson, E. D., et al. 2010, ApJ, 714, 1582, doi: `10.1088/0004-637X/714/2/1582`

Burkert, A. 2015, ApJ, 808, 158, doi: `10.1088/0004-637X/808/2/158`

Cash, W. 1979, ApJ, 228, 939, doi: `10.1086/156922`

Center, S. S. 2015, Spitzer-IRAC Instrument Handbook, edn. SSC, Pasadena, CA

Clark, G. W. 1975, ApJ, 199, L143, doi: `10.1086/181869`

Cohen, J. G. 2004, AJ, 127, 1545, doi: `10.1086/382104`

Cohn, H. N., Lugger, P. M., Couch, S. M., et al. 2010, ApJ, 722, 20, doi: `10.1088/0004-637X/722/1/20`

de Boer, T. J. L., Tolstoy, E., Hill, V., et al. 2012, A&A, 539, A103, doi: `10.1051/0004-6361/201118378`

Degenaar, N., & Wijnands, R. 2009, A&A, 495, 547, doi: `10.1051/0004-6361:200810654`

—. 2010, A&A, 524, A69, doi: `10.1051/0004-6361/201015322`

Dehnen, W., & King, A. 2006, MNRAS: Letters, 367, L29, doi: `10.1111/j.1745-3933.2005.00132.x`

Djorgovski, S. 1993, in Astronomical Society of the Pacific Conference Series, Vol. 50, Structure and Dynamics of Globular Clusters, ed. S. G. Djorgovski & G. Meylan, 373

Dolphin, A. E. 2002, MNRAS, 332, 91, doi: `10.1046/j.1365-8711.2002.05271.x`

Done, C., Gierliński, M., & Kubota, A. 2007, A&A Rev., 15, 1, doi: `10.1007/s00159-007-0006-1`

Engelbracht, C. W., Blaylock, M., Su, K. Y. L., et al. 2007, PASP, 119, 994, doi: `10.1086/521881`

Fabian, A. C., Pringle, J. E., & Rees, M. J. 1975, MNRAS, 172, 15p, doi: `10.1093/mnras/172.1.15P`

Fazio, G. G., et al. 2004, ApJS, 154, 10, doi: `10.1086/422843`

Flesch, E. W. 2015, PASA, 32, e010, doi: `10.1017/pasa.2015.10`

—. 2016, PASA, 33, e052, doi: `10.1017/pasa.2016.44`

Foight, D. R., Güver, T., Özel, F., & Slane, P. O. 2016, ApJ, 826, 66, doi: `10.3847/0004-637X/826/1/66`

Fragos, T., Kalogera, V., Willems, B., et al. 2009, ApJ, 702, L143, doi: `10.1088/0004-637X/702/2/L143`

Frisch, P. 1972, ApJ, 173, 301, doi: `10.1086/151422`

Fruchter, A. S., Levan, A. J., Strolger, L., et al. 2006, Nature, 441, 463, doi: `10.1038/nature04787`

Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, in Proc. SPIE, Vol. 6270, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 62701V

Gilfanov, M. 2004, MNRAS, 349, 146, doi: `10.1111/j.1365-2966.2004.07473.x`

Grindlay, J. E. 1988, in IAU Symposium, Vol. 126, The Harlow-Shapley Symposium on Globular Cluster Systems in Galaxies, ed. J. E. Grindlay & A. G. D. Philip, 347–363

Heinke, C. O. 2010, in American Institute of Physics Conference Series, Vol. 1314, American Institute of Physics Conference Series, ed. V. Kalogera & M. van der Sluys, 135–142

Heinke, C. O., Bahramian, A., Degenaar, N., & Wijnands, R. 2015, MNRAS, 447, 3034, doi: `10.1093/mnras/stu2652`

Heinke, C. O., Grindlay, J. E., Edmonds, P. D., et al. 2005, ApJ, 625, 796, doi: 10.1086/429899

Heinke, C. O., Grindlay, J. E., Lugger, P. M., et al. 2003, ApJ, 598, 501, doi: 10.1086/378885

Heinke, C. O., Wijnands, R., Cohn, H. N., et al. 2006, ApJ, 651, 1098, doi: 10.1086/507884

Hills, J. G. 1991, AJ, 102, 704, doi: 10.1086/115905

Irwin, M., & Hatzidimitriou, D. 1995, MNRAS, 277, 1354, doi: 10.1093/mnras/277.4.1354

Ivanova, N., Heinke, C. O., Rasio, F. A., Belczynski, K., & Fregeau, J. M. 2008, MNRAS, 386, 553, doi: 10.1111/j.1365-2966.2008.13064.x

Ivanova, N., Heinke, C. O., Rasio, F. A., et al. 2006, MNRAS, 372, 1043, doi: 10.1111/j.1365-2966.2006.10876.x

Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS, 399, 683, doi: 10.1111/j.1365-2966.2009.15338.x

Kimmig, B., Seth, A., Ivans, I. I., et al. 2015, AJ, 149, 53, doi: 10.1088/0004-6256/149/2/53

Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102, doi: 10.1088/0004-637X/779/2/102

Kirby, E. N., Guo, M., Zhang, A. J., et al. 2015, ApJ, 801, 125, doi: 10.1088/0004-637X/801/2/125

Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, PASP, 129, 104502, doi: 10.1088/1538-3873/aa80d9

Kundu, A., Maccarone, T. J., Zepf, S. E., & Puzia, T. H. 2003, ApJ, 589, L81, doi: 10.1086/376493

Lasota, J.-P. 2001, New A Rev., 45, 449, doi: 10.1016/S1387-6473(01)00112-9

Liu, J. 2011, ApJS, 192, 10, doi: 10.1088/0067-0049/192/1/10

Maccarone, T. J., Kundu, A., Zepf, S. E., Piro, A. L., & Bildsten, L. 2005, MNRAS: Letters, 364, L61, doi: 10.1111/j.1745-3933.2005.00106.x

Manni, L., Nucita, A. A., De Paolis, F., Testa, V., & Ingrosso, G. 2015, MNRAS. https://arxiv.org/abs/1509.01076

McConnachie, A. W. 2012, AJ, 144, 4, doi: 10.1088/0004-6256/144/1/4

McLaughlin, D. E., & van der Marel, R. P. 2005, ApJS, 161, 304, doi: 10.1086/497429

Monkiewicz, J., Mould, J. R., Gallagher, III, J. S., et al. 1999, PASP, 111, 1392, doi: `10.1086/316447`

Nucita, A., De Paolis, F., Manni, L., & Ingrosso, G. 2013, New A, 23-24, 107, doi: `10.1016/j.newast.2013.03.003`

Perley, D. A., Quimby, R. M., Yan, L., et al. 2016, ApJ, 830, 13, doi: `10.3847/0004-637X/830/1/13`

Perlman, E. S., Padovani, P., Giommi, P., et al. 1998, AJ, 115, 1253, doi: `10.1086/300283`

Pietrukowicz, P. 2009, Acta Astron., 59, 291. `https://arxiv.org/abs/0910.1302`

Piro, A. L., & Bildsten, L. 2002, ApJ, 571, L103, doi: `10.1086/341341`

Podsiadlowski, P., Pfahl, E., & Rappaport, S. 2005, Binary Radio Pulsars, 328

Podsiadlowski, P., Rappaport, S., & Pfahl, E. D. 2002, ApJ, 565, 1107, doi: `10.1086/324686`

Pooley, D., Lewin, W. H. G., Anderson, S. F., et al. 2003, ApJ, 591, L131, doi: `10.1086/377074`

Psaltis, D. 2004, ArXiv Astrophysics e-prints

Ramsay, G., & Wu, K. 2006, A&A, 459, 777, doi: `10.1051/0004-6361:20065959`

Regis, M., Richter, L., Colafrancesco, S., et al. 2015, MNRAS, 448, 3731, doi: `10.1093/mnras/stu2747`

Revnivtsev, M., Postnov, K., Kuranov, A., & Ritter, H. 2011, A&A, 526, A94, doi: `10.1051/0004-6361/201014726`

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, ApJS, 166, 470, doi: `10.1086/506525`

Rieke, G. H., et al. 2004, ApJS, 154, 25, doi: `10.1086/422717`

Saeedi, S., Sasaki, M., & Ducci, L. 2016, A&A, 586, A64, doi: `10.1051/0004-6361/201526233`

Sbordone, L., Bonifacio, P., Buonanno, R., et al. 2007, A&A, 465, 815, doi: `10.1051/0004-6361:20066385`

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103, doi: `10.1088/0004-637X/737/2/103`

Schuster, M. T., Marengo, M., & Patten, B. M. 2006, in Proc. SPIE, Vol. 6270, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 627020

Schweitzer, A. E., Cudworth, K. M., Majewski, S. R., & Suntzeff, N. B. 1995, AJ, 110, 2747, doi: `10.1086/117727`

Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48, doi: `10.1088/0004-637X/788/1/48`

Shaw, A. W., Heinke, C. O., Degenaar, N., et al. 2017, MNRAS, 471, 2508, doi: `10.1093/mnras/stx1732`

Siegel, M. H., Dotter, A., Majewski, S. R., et al. 2007, ApJ, 667, L57, doi: `10.1086/522003`

Sonbas, E., Rangelov, B., Kargaltsev, O., et al. 2016, ApJ, 821, 54, doi: `10.3847/0004-637X/821/1/54`

Spitzer, Jr., L., Cochran, W. D., & Hirshfeld, A. 1974, ApJS, 28, 373, doi: `10.1086/190323`

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163, doi: `10.1086/432523`

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Tendulkar, S. P., Bassa, C. G., Cordes, J. M., et al. 2017, ApJ, 834, L7, doi: `10.3847/2041-8213/834/2/L7`

Tinney, C. G. 1999, MNRAS, 303, 565, doi: `10.1046/j.1365-8711.1999.02233.x`

Tinney, C. G., Da Costa, G. S., & Zinnecker, H. 1997, MNRAS, 285, 111, doi: `10.1093/mnras/285.1.111`

Tolstoy, E., Venn, K. A., Shetrone, M., et al. 2003, AJ, 125, 707, doi: `10.1086/345967`

van Paradijs, J., & McClintock, J. E. 1995, X-ray Binaries, 58

Verbunt, F., & Freire, P. C. C. 2014, A&A, 561, A11, doi: `10.1051/0004-6361/201321177`

Verbunt, F., & Hut, P. 1987, in IAU Symposium, Vol. 125, The Origin and Evolution of Neutron Stars, ed. D. J. Helfand & J.-H. Huang, 187

Verbunt, F., & Lewin, W. H. G. 2006, Globular cluster X-ray sources, ed. W. H. G. Lewin & M. van der Klis (Cambridge University Press), 341–379

Verbunt, F., van den Heuvel, E. P. J., van Paradijs, J., & Rappaport, S. A. 1987, Nature, 329, 312, doi: `10.1038/329312a0`

Véron-Cetty, M.-P., & Véron, P. 2010, A&A, 518, A10, doi: `10.1051/0004-6361/201014188`

Walker, M. G., Mateo, M., & Olszewski, E. W. 2009, AJ, 137, 3100, doi: `10.1088/0004-6256/137/2/3100`

Wang, S., Liu, J., Qiu, Y., et al. 2016, ApJS, 224, 40, doi: `10.3847/0067-0049/224/2/40`

White, N. E., & Ghosh, P. 1998, ApJ, 504, L31, doi: `10.1086/311568`

White, III, R. E., Sarazin, C. L., & Kulkarni, S. R. 2002, ApJ, 571, L23, doi: `10.1086/341122`

Chapter 3

Identifying New X-ray Binary Candidates in M31 using Random Forest Classification

3.1 Introduction

Galaxies are key ingredients in understanding the evolution of the Universe. They contain the majority of the stars, gas, dust, planets, and other objects of interest. Galaxies and their contents are the majority of the luminous matter in the universe.

However, studying our own Milky Way Galaxy on a large scale is often challenging. Our position within the Galaxy is not favourable for studying its structure, since dust and gas in the disk obscures our view along important lines of sight. Nearby galaxies are sufficiently close that we can resolve their structure at a favourable viewing angle, while still being smaller in angular size than the Milky Way. Meanwhile, more distant galaxies at higher redshift can tell us about the past evolution of the Milky Way, and more generally trace the evolution of galaxies at different epochs. Studying bulk properties of galaxies gives us clues about how processes happen on scales that we may not be able to resolve for an individual galaxy. One process which can be resolved in nearby galaxies is the emission of X-ray binaries (XRBs).

3.1.1 X-ray Binaries and Galaxies

Aside from possible diffuse emission and the presence of an active galactic nucleus (AGN), the X-ray emission of nearby galaxies is dominated by the presence of X-ray binaries, relatively rare systems that contain a compact object in a close binary with an ordinary star. All X-ray binaries can be categorized by the type of compact object accreting material from the companion: those with black hole (BH) and those with neutron star (NS) primaries. The majority of X-ray binaries can also be classified based on the type of the companion star. Compact objects which accrete from a $< 1\ M_\odot$ companion undergoing Roche Lobe overflow are known as low-mass X-ray binaries (LMXBs), while compact objects accreting from a $\geq 10\ M_\odot$ star, usually through the stellar wind, are identified as high-mass X-ray binaries (HMXBs) (Casares et al., 2017).

The advent of high spatial resolution, sensitive X-ray observations in the *Chandra* era has permitted the study of X-ray sources in nearby galaxies on resolved scales. Aside from their value as laboratories for extreme physics, XRBs can be used as tracers of galaxy properties. The X-ray luminosity functions (XLFs) of sources within nearby star-forming galaxies are dominated by contributions from HMXBs. The luminosity functions of different galaxies can occupy a variety of locations in the $N(< L) - L$ space. However, if each XLF is normalized by the galaxy's star formation rate, then galaxies tend to have similar XLFs occupying a narrow band in $N(< L) - L$ space (Grimm et al., 2003). This trend appears consistent at resolved scales: in the Milky Way, HMXBs cluster spatially close to known active star-forming complexes in the Milky Way's spiral arms (Bodaghee et al., 2012).

LMXBs also provide a useful proxy measurement of galaxy properties. Many LMXBs are found in the globular clusters of galaxies where they can be created through dynamical encounters enabled by the high stellar densities (Verbunt & Lewin, 2006). Consequently, the fraction of a galaxy's LMXBs that are found in globular clusters increases with the specific frequency of globular clusters (Maccarone et al., 2003). In addition, low mass stars comprise the bulk of the stellar mass in any stellar population, so LMXBs within a galaxy can trace their host galaxy's stellar mass (Gilfanov, 2004). LMXBs can be viewed, therefore, as stellar mass and stellar encounter rate tracers.

3.1.2 Identifying X-Ray Binaries

To use XRBs as a probe for galaxy structure and properties, accurate determinations of the XRB count in a population are required. In principle, classifying X-ray sources based only on their X-ray emission is difficult. Depending on the energy bands used for analysis of an observation, supernova remnants in the same galaxy can mimic the X-ray appearance of X-ray binaries (Grimm et al., 2003). Nearby galaxies, with a large angular size on the sky, are more likely to have X-ray source lists contaminated by X-ray active foreground stars in the Milky Way or background AGN. The preferred method to resolve this source confusion is to identify longer-wavelength counterparts, as spectroscopy can separate AGN from expected XRB counterparts. In addition, sources that are spatially extended at longer wavelengths can be identified as associated with distant galaxies and therefore likely AGN, while sources that have high proper motion or an extreme optical to X-ray flux ratio can be identified as foreground stars (Vilhu & Walter, 1987; Guillot et al., 2009; Saeedi et al., 2016).

Multiwavelength observations of X-ray binary candidates may not be available or practical due to crowding, extinction, or large distance. In the absence of such multiwavelength observations, classification tends to rely on the presence of X-ray features unique to compact objects (or neutron stars in particular), such as Type I X-ray bursts or X-ray pulsations. If these X-ray features are also not present, classification is done using a combination of X-ray brightness and the X-ray colour or hardness ratio. X-ray colour-colour diagrams or colour-colour-intensity diagrams show that certain types of objects tend to cluster together based on, for example, compact object type and pulsating versus non-pulsating neutron stars (Prestwich et al., 2003; Vrtilek & Boroson, 2013). However, this method tends to result in approximate decision boundaries that are difficult to constrain. The launch of the *NuSTAR* X-ray telescope has allowed images of nearby galaxies in hard X-rays ($>$ 10 keV) to be constructed for the first time (Wik et al., 2014; Yukita et al., 2016). Recent work using hardness-intensity and colour-colour diagrams constructed from observations of nearby galaxies have suggested that the properties of X-ray sources at energies greater than 10 keV can help separate some X-ray source types, for example black hole versus neutron star binaries (Vulic et al., 2018). However, the sensitivity of hard X-ray telescopes is such that only the brightest sources are detected, therefore catalogues produced by observations with hard X-ray telescopes have far fewer sources than those produced from observations at softer X-ray energies, such as with *Chandra*.

3.1.3 Machine Learning

One approach to overcoming the approximate decision boundaries found through simple colour-colour or hardness-intensity diagrams is to include more information from available spectral information and the use of machine learning. Supervised machine learning (ML) describes a series of techniques where a function/algorithm "learns" outputs based on a previously trained set of paired input/output data. These techniques can predict either as a regression or a classification when applied to new, unknown examples. Machine learning classifiers tend to perform optimally on large datasets with many classified examples and a sufficient number of informative features which can help define a model that discriminates between different desired categories. In astronomy, machine learning has already been used to investigate classification problems where the data may have high dimensionality and is difficult to either model or assign simple classification boundaries. Recently, Ksoll et al. (2018) used machine learning techniques to separate lower main sequence stars from pre-main sequence stars using a random forest algorithm applied to the Hubble Space Telescope (HST) photometry of 30 Doradus. Machine learning has previously been used to investigate the properties of X-ray binaries. Sonbas et al. (2015) used a learning decision tree algorithm to classify X-ray sources in the Draco dwarf galaxy, finding that classifications made by that method tend to agree with classifications made with previously established classification techniques (i.e., identification of multiwavelength counterparts; Brehm et al. 2014). Gopalan et al. (2015) expanded the colour-colour-intensity diagram classification technique by applying a supervised learning algorithm as a method of demarcating systems containing black holes, pulsating neutron stars, or non-pulsating neutron stars.

3.1.4 Andromeda Galaxy

Nearby galaxies are excellent places to test X-ray source classification techniques, as sources in their direction can be simplified to either be sources inside the galaxy at its distance, foreground stars, or distant background galaxies. M31, the Andromeda galaxy, is the nearest large galaxy to our own Milky Way, and the closest spiral galaxy. Although more massive and containing a larger stellar population, M31 is similar to the Milky Way and is viewed at a relatively favourable viewing angle without many of the observational complications inherent to observing the Milky Way itself. M31 has been catalogued in X-rays to a luminosity limit of 10^{34} erg s^{-1} in three distinct fields. These fields are located in the bulge and two portions of the disk (southwest and northwest). In total, over 900 X-ray sources have been identified across these three fields. This means that M31 represents the best test case for machine learning approaches because ML algorithms typically perform much better with higher numbers of classified (and total) examples. This, along with the abundance of multiwavelength data available, such as the

HST Panchromatic Hubble Andromeda Treasury, makes M31 ideal for testing machine learning techniques since we can use that multiwavelength data to evaluate the efficacy of machine learning algorithms at identifying new XRB candidates.

One issue with studying the XRB content of nearby galaxies is that their large angular sizes on the sky means that contamination by foreground stars and background AGN is significant. For M31, the background AGN population has been extensively investigated compared to other nearby galaxies (e.g. Dorn-Wallenstein et al., 2017; Huo et al., 2015; Meusinger et al., 2010). Our goal is to improve on previous investigations of the nature of X-ray sources in nearby galaxies by considering a large dataset of *Chandra*-detected X-ray sources in the field of M31 and testing multiple supervised machine learning algorithms at once. This will hopefully provide clues about how to develop a better method of classifying X-ray sources using only their X-ray emission in the absence of other discriminating data.

3.2 Data and Method

3.2.1 *Chandra* Data

As our sample dataset, we consider the catalogue of *Chandra* X-ray sources in M31 from Vulic et al. (2016). A full description of the catalogue can be found in that paper, however we summarize the key characteristics here. This catalogue resulted from combining all 133 available ACIS-I/S observations of M31 to detect sources in three distinct fields (bulge, northwest, and southwest) at a 0.5–8.0 keV luminosity limit of 10^{34} erg s^{-1}, covering a total area of ~ 0.6 deg^2.

In total, the catalogue consists of 943 sources. There are more sources than reported in Vulic et al. (2016) because we also include sources that did not meet the "probability of no-source" cutoff of 1×10^{-2} used as a filter in that study. Each source has, in addition to a variety of observational features, the position, median incident energy, mean observed energy, and the photon flux in 16 different energy bands between 0.5–8.0 keV. These energy bands, the `ACIS EXTRACT` defaults, can often be represented as linear combinations of other energy bands, and as such they are likely to be highly correlated. A source only needs to be detected in one of the 16 energy bands to be part of the catalogue, and for most sources at least one of the bands will have a flux that is zero or consistent with zero within error.

3.2.2 HST PHAT Cluster Data

Aside from internal performance metrics, we attempted to validate our technique by comparing classifications from our best-performing algorithms to a classification not based on X-ray source properties. For this, we used data from the Panchromatic Hubble Andromeda Treasury (PHAT), a large multi-year survey that obtained *Hubble Space Telescope* photometry for roughly a third of M31's disk in multiple filters across 23 "bricks" of observations (see Dalcanton et al. 2012; Williams et al. 2014 and accompanying articles). This survey permits some 100 million individual stars and other objects of interest to be resolved. We use the star cluster and background galaxy catalogs of Johnson et al. (2015), which provides a library of 2,753 clusters and 2,270 background galaxies in the field of M31. These clusters and background galaxies were catalogued from a citizen-science project that classified objects based on their morphology in PHAT images. We use the PHAT catalogs of clusters, background galaxies, and unknown sources given by Johnson et al. (2015). Only a few sections of the *Chandra* data have PHAT coverage, since PHAT only covers the northeast disk of M31; the coverage of both the *Chandra* data and the PHAT footprint are shown in Figure 3.1.

3.2.3 Feature Construction

In order to evaluate the dataset using machine learning techniques, we must first construct feature inputs for the algorithms. In order to construct distance-independent features, we divide each of the fluxes by the broad band 0.5–8.0 keV flux. We therefore use, for each energy band, the ratio of the flux emitted in that particular band to the total flux measured. Our final list of features consists of the fifteen ratios, the 0.5–8.0 keV flux, and the median incident and observed energy of each source. The number of sources which have a non-zero value for each feature in both the classified and unclassified sets is given in Table 3.1. Since uncertainties may result in measured photon fluxes that are negative or that exceed the overall 0.5–8.0 keV photon flux, it is possible that computing flux ratios may yield aphysical values that are negative or greater than 1. Crowding near the bulge may make 0.5–8.0 keV fluxes unexpectedly smaller because of contamination from nearby sources - `ACIS EXTRACT` handles crowding by shrinking extraction regions. Since they contain all of the available counts, the 0.5–8.0 keV flux is more likely to be affected by this crowding than other bands. For physical reasons, we set the photon flux ratios to be zero if they are less than zero, and to one if they are greater than one.

3.2.4 Classification Scheme

We create a training set by using sources that are classified by Vulic et al. (2016) through crossmatching with sources classified in the literature, principally the *XMM-Newton* source catalogue of Stiele et al. (2011). In total, there are 163 previously classified sources, of which 77 are X-ray binaries (XRB), 43 are background active galactic nuclei (AGN), 29 are foreground stars (fgStar), and 14 are supernova remnants (SNR). Since we are primarily interested in the identification of new XRB candidates, we use machine learning algorithms to classify the unknown objects in two ways: first, we consider a multiclass classification where we attempt to classify new objects as XRB, AGN, fgStar, or SNR. Secondly, we consider a binary classification where we attempt to determine if an object is an XRB or not. The multiclass classification allows us to evaluate the viability of classification across multiple object types, while the binary classification allows us to use ML performance metrics (e.g., receiver-operating characteristic curves) that require a two-class formulation of the problem. Additionally, given the small number of classified sources overall (< 200), we can expect that a classification scheme with only two categories will perform better. The binary classification scheme also has a more even distribution of objects between classes, compared to the multiclass scheme.

3.2.5 Algorithms

To explore the dataset using machine learning algorithms, we use supervised learning algorithms from the Python `sklearn` package, version 0.19.1 (Pedregosa et al., 2011). These

Table 3.1: Summary of dataset properties

Feature Name	# classified objects	# unclassified objects
0.5 – 8.0 keV photon flux	163	780
0.5 – 2.0 keV photon flux fraction	163	749
2.0 – 8.0 keV photon flux fraction	153	744
0.5 – 1.7 keV photon flux fraction	163	736
1.7 – 2.8 keV photon flux fraction	152	679
2.8 – 8.0 keV photon flux fraction	147	723
0.5 – 1.5 keV photon flux fraction	162	728
1.5 – 2.5 keV photon flux fraction	156	684
2.5 – 8.0 keV photon flux fraction	149	731
0.5 – 1.0 keV photon flux fraction	155	634
1.0 – 2.0 keV photon flux fraction	163	719
2.0 – 4.0 keV photon flux fraction	148	686
4.0 – 6.0 keV photon flux fraction	139	636
6.0 – 8.0 keV photon flux fraction	121	513
0.5 – 7.0 keV photon flux fraction	163	779
2.0 – 7.0 keV photon flux fraction	152	742
Mean Observed Energy	163	768
Mean Incident Energy	156	664

The number of classified and unclassified objects per feature varies because some objects have feature values set to zero due to a negative flux or energy being inferred from ACIS EXTRACT.

algorithms use a set of already-classified training data as input for classifying new data. Since training and evaluation of a dataset of this size is relatively quick, we evaluated multiple algorithms at once. Each algorithm used has a number of initialization parameters, also known as “hyperparameters,” which change the fitting behaviour of the algorithm. We evaluated each of these hyperparameters by performing a one-dimensional search over gridded values of the hyperparameter to look for the best value. The scoring for the best value is specified by the user – in this case we used the cross-validation score (see Section 3.3 for an explanation) as the determinant of the best estimator. In many cases, there was no obvious trend for a “best” value for a given hyperparameter. In cases where there was a clear “best” value of the hyperparameter, we use (and specify) that value. Otherwise, we use the default value of that hyperparameter. We tested multinomial logistic regression, Gaussian naive Bayes, random forest, a linear support vector classifier, and a multi-layer perceptron neural network. We chose to evaluate these algorithms since they are commonly-used machine learning algorithms for a variety of classification tasks.

Logistic regression is a generalized model-fitting technique similar to linear regression, except that it attempts to fit to probability of class membership instead. In general, it assumes that class membership is linearly separable in the feature space. Based on results from simpler techniques of classifying X-ray sources, such as hardness ratio diagnostics, we do not expect that our categories of X-ray source are linearly separable in the feature space. However, logistic regression provides a useful baseline comparison and could be considered similar to a simple classification cut made in the feature space. In our logistic regression model, we used the following `sklearn` hyperparameters on the `sklearn.linear_model.LogisticRegression()` function: balanced class weights, one-versus-rest multi-class handling, L1 penalty with a SAGA (Stochastic Average Gradient with support for L1 regularization) solver, inverse regularization strength = 1.0 and a stopping tolerance of 0.001.

Gaussian naive Bayes is a model that produces conditional class probabilities using a Bayesian formulation with the additional assumption that all features are conditionally independent from each other given the class label. Since our feature set is not conditionally independent in general, it also provides a useful baseline for comparison. We trained the naive Bayes algorithm `sklearn.naive_bayes.GaussianNB()` using default parameters.

The random forest method uses the aggregate results of an ensemble of decision trees that have been fit on a subset of features and samples. Each of these decision trees uses a loss function to divide up the samples by segmenting the feature space until all of the “leaves” contain samples of only one type. In this case, the loss function is a function that optimizes feature space segmentations (branches of the decision tree) to have samples of only one type with the fewest number of segmentations. The random forest classifies new samples as the classification returned from the majority of the decision trees in the forest.

Although individual decision trees are highly biased towards the subsample of data/features they fit from, in aggregate the random forest is not strongly biased by its training set. In addition, random forest algorithms are typically useful even in cases when features are not normalized and when there are a relatively small number of features in the dataset. We trained the random forest `sklearn.ensemble.RandomForestClassifier()` using 500 decision trees, with balanced class weights, made splits using an entropy/information gain loss function, and with a maximum tree depth of 80.

Linear support vector classification (SVC) is a technique which fits a separating hyperplane

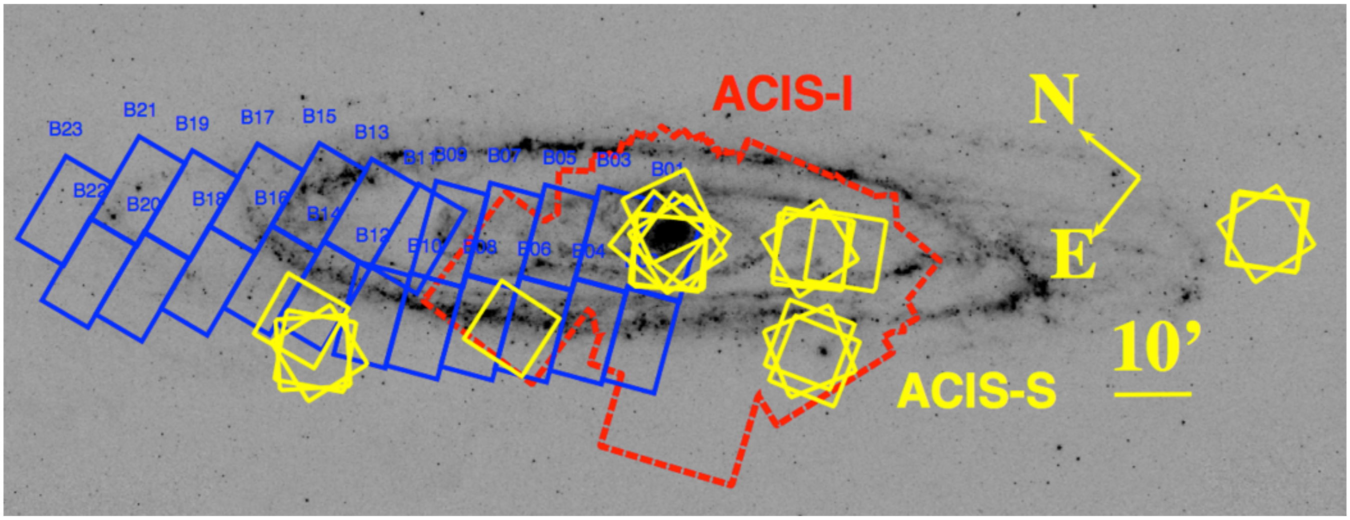

Figure 3.1: Locations of *Chandra* ACIS-I (red dashed outline) and ACIS-S3 merged observations (yellow solid lines) used in our analysis (see Vulic et al., 2016), with PHAT footprint (blue rectangles). Background image: *Spitzer*/MIPS 24 μm observations (Gordon et al., 2006).

in the feature space that can be used for classification of future examples. We fit the `sklearn` linear SVC `sklearn.svm.LinearSVC()` using default parameters.

Finally, we used a multi-layer perceptron, which is a class of neural network that learns a non-linear function to classify samples. It possesses non-linear hidden layers that learn between the feature inputs and the fitted output. Multi-layer perceptrons are advantageous in that they learn non-linear functions well, but they often require extensive hyperparameter tuning to be effective. We used a multi-layer perceptron `sklearn.neural_network.MLPClassifier` with 1 hidden layer of 100 neurons, logistic activation, an LBFGS (limited-memory Brodyen-Fletcher-Goldfarb-Schanno algorithm) solver, and a constant learning rate initialized at 0.01.

We also tested the random forest algorithm available through the **R** randomForest package v. 4.6-12 (Liaw & Wiener, 2002; R Core Team, 2018). Given that we find the random forest to have the overall best performance of the sklearn algorithms (see below in Sections 3.3 and 3.4), we wish to compare the random forest implementations from two of the most popular machine learning packages, expecting that they should give similar performance. We used identical hyperparameters to the `sklearn` random forest (where such hyperparameters could be specified) in order to compare similar realizations of the fitted algorithm.

3.3 Multiclass Results

We evaluated the algorithms in multiple ways. Firstly, we randomly split the classified samples 70%/30% into a training and test set, training each algorithm on the majority of the samples and testing them on the remainder. This training/test split is a relatively common ratio in machine learning problems chosen to avoid overfitting (e.g., Ksoll et al., 2018). The accuracy, defined as the number of correct classifications divided by the total number of classifications, was computed on one particular realization of the training/test split. In addition, we computed

Table 3.2: Algorithm Evaluation, multiclass case

Algorithm	Accuracy	Recall	CV Score
Logistic Regression	0.55	0.55	0.54 ± 0.04
Naive Bayes	0.57	0.57	0.52 ± 0.07
Support Vector Classification	0.49	0.49	0.55 ± 0.04
Random Forest (sklearn)	0.57	0.57	0.65 ± 0.06
Multi-layer Perceptron NN	0.57	0.57	0.52 ± 0.08
Random Forest (R)	0.61	0.61	0.66 ± 0.07

the recall on this realization of the training/test split, which is defined per class (e.g., XRB, SNR, AGN, fgStar) as the number of correct classifications of that class divided by the total number of true members of that class. Since we have significant class imbalance, we computed the recall as the average recall across each class weighted by the number of true instances of that class. The accuracy and recall of the trained model as applied to the test set is tabulated in Table 3.2. We also computed confusion matrices, which track the predicted versus actual class for all objects in the sample set. Secondly, we performed k-fold cross-validation on the entire classified dataset. In k-fold cross-validation, the dataset is partitioned into k subsamples. Each of the k subsamples is used as validation for a model trained on the remaining $k-1$ subsamples, and the cross-validation score is calculated as the mean and standard deviation of the accuracy across each of the k trials. In our cross-validation, we chose $k = 5$, as it is typically chosen as an intermediate value between high values (which yield excessively high bias) and low values (which yield excessively high variance).

As can be seen from the results in Table 3.2, performance metrics for the algorithms are generally poor in the multiclass scenario. However, based on the metrics (accuracy, recall, cross-validation score), the random forest algorithms (both sklearn and R) tend to have the best performance, and the multi-layer perceptron neural network the poorest. Naively, if the classes were balanced, we would expect that the accuracy should be ~ 0.25 for guessing randomly. However, we have significant class imbalance, with far more XRBs than any other class. A randomly guessing algorithm would tend to always pick the most numerous class, which would give an accuracy of $77/163 \approx 0.47$. Our algorithms predict better than randomly guessing, however this is not a sufficient baseline for evaluating performance, as class imbalance means that models which always predict the majority class (or classes) will perform well. The main source of poor performance is not for classifications of true X-ray binaries. In the confusion matrices for the random forest algorithms, shown in Tables 3.3 and 3.4, the accuracy and recall is primarily being reduced by mis-classifications of objects other than XRBs, especially foreground stars and supernova remnants, which are less numerous. The algorithm also has difficulty classifying AGN, which are classified roughly evenly into the three other categories.

3.4 Two-class Results

We also consider re-evaluation of the problem as a binary one, where we reassign each object to be defined as either an XRB or a non-XRB. Since identifying new XRB candidates is the primary goal of this ML-based classification, this permits us to pursue a method that is poten-

Table 3.3: Confusion matrix for sklearn, multiclass case

Predicted Class \ Actual Class	AGN	SNR	fgStar	XRB	Total
AGN	5	0	2	2	9
SNR	0	3	1	0	4
fgStar	3	2	5	0	10
XRB	6	1	4	15	26
Total	14	6	12	17	49

Table 3.4: Confusion matrix for R, multiclass case

Predicted Class \ Actual Class	AGN	SNR	fgStar	XRB	Total
AGN	6	1	3	2	12
SNR	0	4	1	0	5
fgStar	3	1	6	0	10
XRB	5	0	2	15	22
Total	14	6	12	17	49

tially more accurate overall and is not limited by strong class imbalance. In addition, using a binary realization of the problem permits us to evaluate algorithm performance using metrics that can only be applied to binary problems, such as the area under the curve (AUC) of a receiver-operating curve.

As with the multiclass case, we performed both a 70/30 training/test split to the classified samples and used the predictions to generate confusion matrices, accuracy, and recall measures. We also performed 5-fold cross-validation on the classified samples and compute the CV score as the mean accuracy across each of the 5 folds. The scores for each metric are tabulated in Table 3.5. In addition, we also create receiver-operating characteristic (ROC) curves for the sklearn algorithms. This curve plots the true positive rate against the false positive rate at various classification thresholds for the binary classifier (Spackman, 1989). The area under this curve (AUC) is interpreted as the probability that the classifier will rank a randomly chosen true positive example higher than a randomly chosen true negative example. As such, a classifier which guesses randomly (or is uninformative) would have AUC = 0.5. We plot ROC curves for the sklearn algorithms in Figure 3.6, and tabulate the AUCs in Table 3.5. Each of the algorithms gives an AUC significantly better than an uninformative classifier, though once again the random forest implementations (both sklearn and R) tend to give the best overall performance. The confusion matrices for the random forest algorithms are shown in Tables 3.6 and 3.7. In contrast with the multiclass approach, the **R** and sklearn algorithms return identical confusion matrices. This is most likely due to the binary case being less noisy and having less complex decision trees. These confusion matrices also illustrate the gains in accuracy made by grouping all non-XRB categories together - the number of true and false positive XRB sources is similar to the multiclass case. The overall number of incorrect classifications is reduced compared with the multiclass case where the majority of misclassifications were amongst the three non-XRB classes, which are now correctly classified as simply non-XRB sources.

Table 3.5: Algorithm Evaluation, binary case

Algorithm	Accuracy	Recall	AUC*	CV Score
Logistic Regression	0.71	0.77	0.74	0.66 ± 0.06
Naive Bayes	0.73	0.77	0.85	0.74 ± 0.09
Support Vector Class.	0.71	0.77	0.85	0.71 ± 0.08
Random Forest (sklearn)	0.84	0.85	0.88	0.75 ± 0.05
Multi-layer Perceptron	0.59	0.63	0.63	0.72 ± 0.07
Random Forest (R)	0.86	0.86	0.89	0.79 ± 0.06

*Area Under Curve

Table 3.6: Confusion matrix for sklearn, binary case

		Actual Class		
		XRB	non-XRB	Total
Predicted Class	XRB	15	5	20
	non-XRB	2	27	29
	Total	17	32	49

Table 3.7: Confusion matrix for R, binary case

		Actual Class		
		XRB	non-XRB	Total
Predicted Class	XRB	15	5	20
	non-XRB	2	27	29
	Total	17	32	49

Table 3.8: PHAT Crossmatches to Unidentified X-ray Sources

Source ID	RA	Dec	RF Class multi	P_{XRB}	RF Class binary	P_{XRB}	PHAT Class
004325.64+411537.4	10.856839	41.260392	AGN	0.29	no	0.19	cluster
004233.25+411742.2	10.638557	41.295061	XRB	0.85	yes	0.84	cluster
004250.81+411707.3	10.711741	41.285387	XRB	0.56	yes	0.57	cluster
004255.60+411835.0	10.731686	41.309739	XRB	0.74	yes	0.74	cluster
004246.08+411736.1	10.692008	41.293385	XRB	0.79	yes	0.78	cluster
004615.36+414128.1	11.564025	41.691153	AGN	0.33	no	0.3	unknown
004614.67+414317.6	11.561131	41.721574	AGN	0.32	no	0.31	unknown
004616.82+414300.4	11.570103	41.716791	AGN	0.17	no	0.1	unknown
004248.83+411512.9	10.703484	41.253587	fgStar	0.0	no	0.03	unknown

3.5 PHAT Crossmatching

To provide an independent method of evaluating our algorithms' classification strength that is not based on X-ray properties, we crossmatched the *Chandra* catalogue of M31 sources to clusters, background galaxies, and unclassified sources found in the PHAT survey by Johnson et al. (2015). We used a 0.5" tolerance between the two catalogues, finding 37 matches. If an X-ray source is co-located with a PHAT-identified cluster or background galaxy, it is overwhelmingly likely that that X-ray source is an X-ray binary or AGN, respectively. The majority of these matches are XRBs or background galaxies which are already identified as such by previous surveys. However, we identify nine sources with a PHAT counterpart that are not identified in previous surveys, whose properties we tabulate in Table 3.8. Five of these objects are morphologically classified as clusters, while four are not identified either as star clusters or background galaxies. Of the four sources with PHAT counterparts that can't be identified, both our binary and multiclass algorithms classify these objects as non-XRBs—the multiclass algorithm classifies three as AGN and one as a foreground star. Of the five sources associated with clusters, our algorithm classifies four as XRBs and one as a background AGN. This 80% accuracy of classifying XRBs is in line with our expectations from the binary results, as discussed above in Section 3.4, though we caution that our number of matched sources is very small and this sample is not statistically significant.

3.6 Discussion

We find that algorithms perform significantly better using a binary approach (XRB vs non-XRB) rather than a multiclass approach. This is expected for two reasons. Firstly, we have fairly significant class imbalance and a low number of classified samples overall. XRBs are the most numerous class and only a handful of SNRs and fgStars are present by comparison. In general, the performance of an algorithm depends strongly on the number of available classified examples; the precise number necessary depends on the structure of the feature space and the desired significance levels (Raudys & Jain, 1991). Secondly, classification using only X-ray emission in the narrow energy range of telescopes like *Chandra* (e.g., 0.5–8.0 keV) is expected

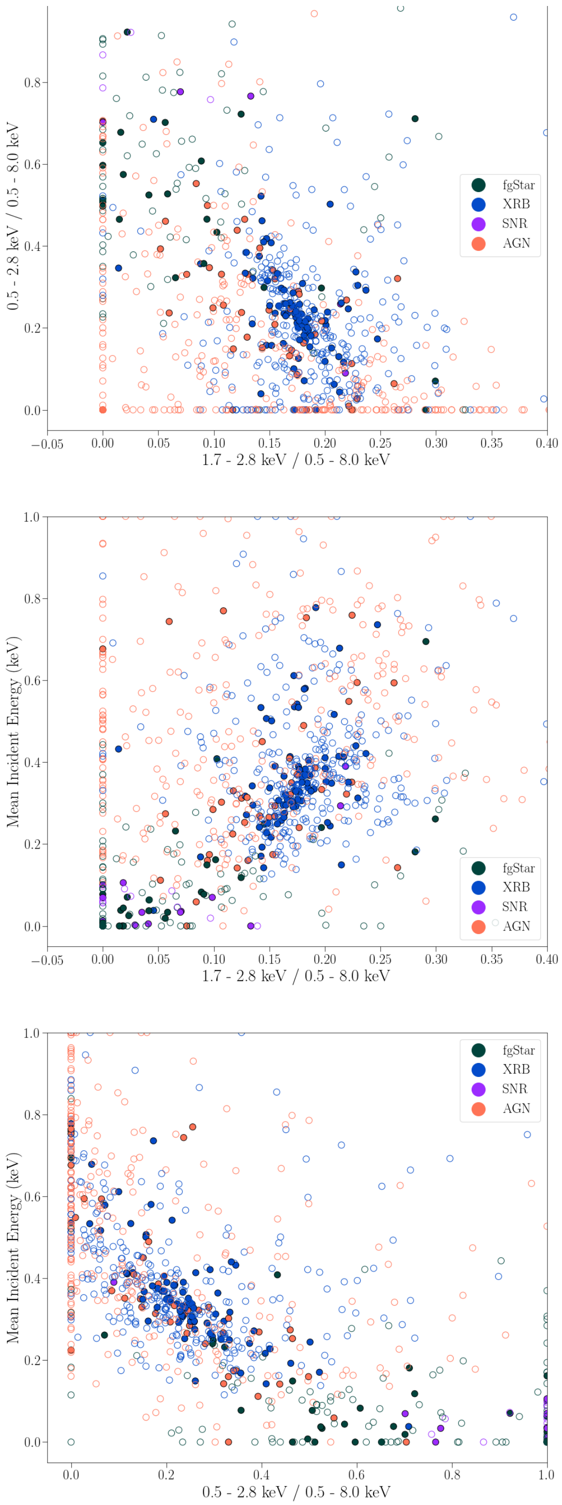

Figure 3.2: Feature space for three of the most significant features in the multiclass approach as determined by the `sklearn` random forest. Since a number of the features determined to be most significant are similar, we chose to plot the features that are the most distinct yet still have high significance. The piling up of features at 0.0 and 1.0 on each plot is due to the tacking of values outside of this range to these boundaries in order for the photon flux fractions to be physically interpretable.

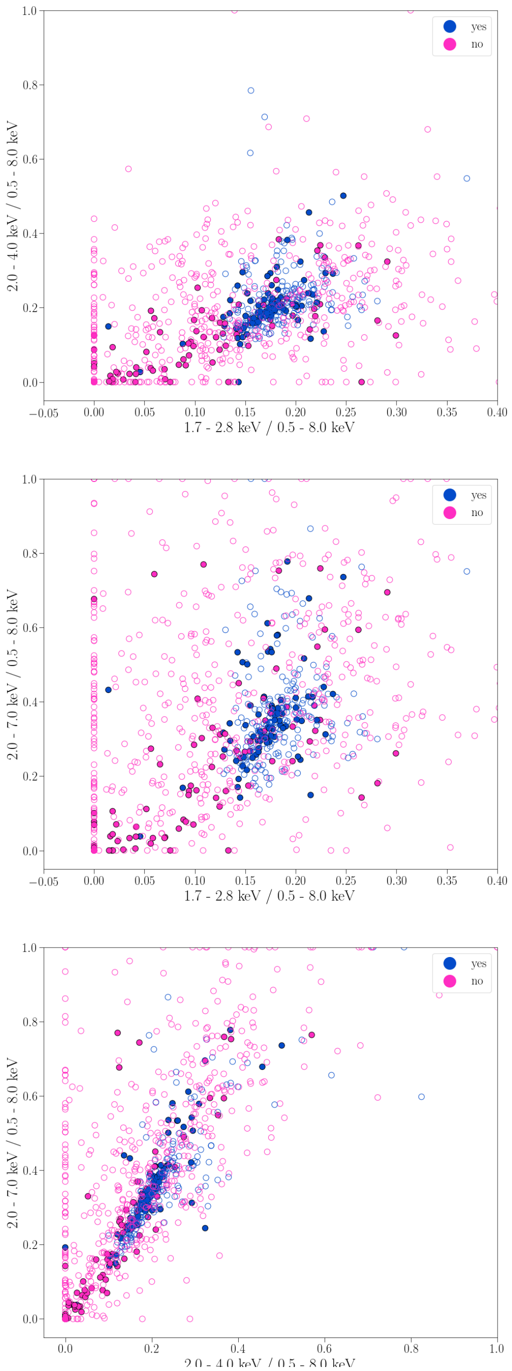

Figure 3.3: Feature space for three of the most significant features in the binary approach as determined by the `sklearn` random forest. Since a number of the features determined to be most significant are similar, we chose to plot the features that are the most distinct yet still have high significance. The piling up of features at 0.0 and 1.0 on each plot is due to the tacking of values outside of this range to these boundaries in order for the photon flux fractions to be physically interpretable.

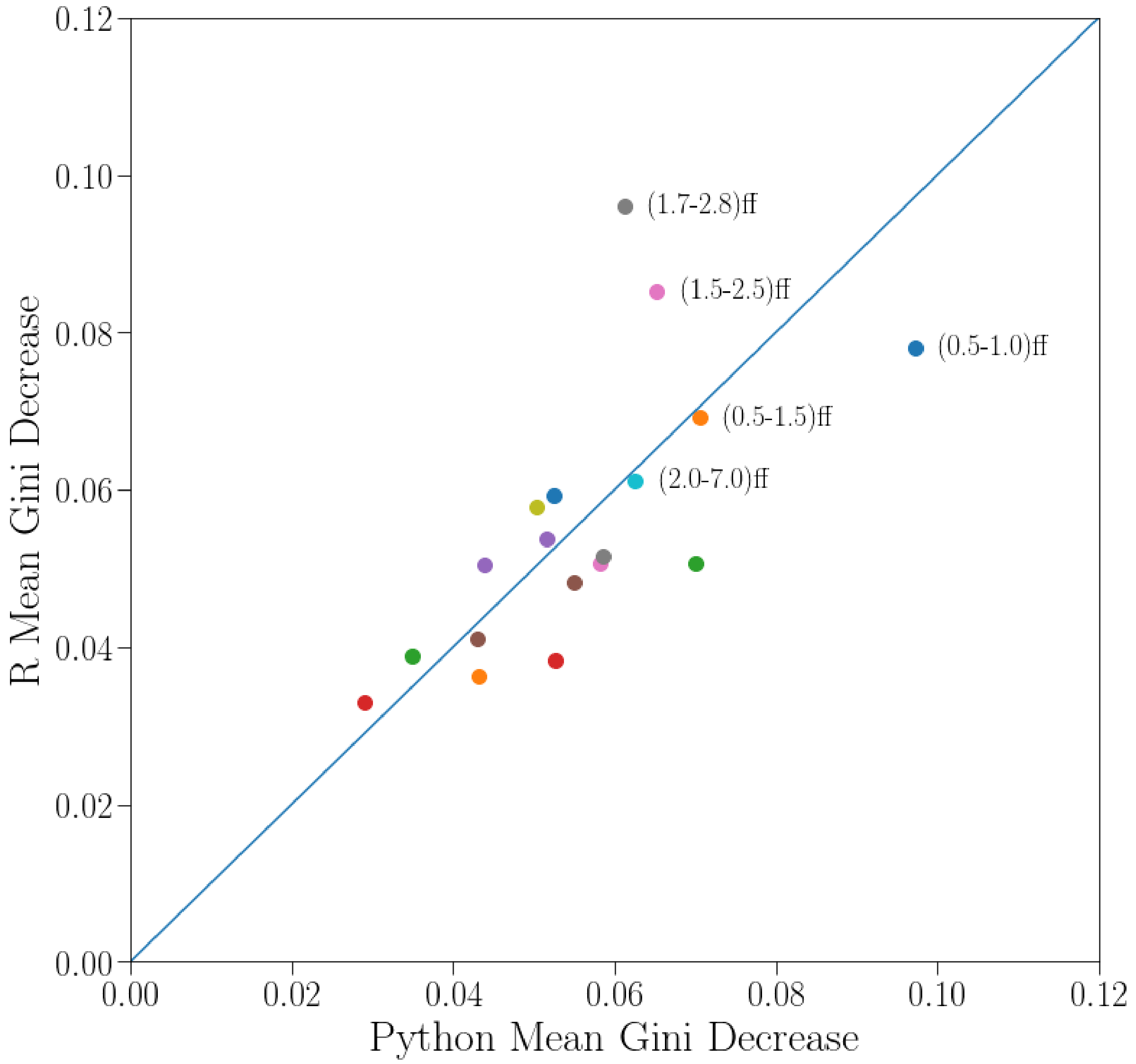

Figure 3.4: Comparison of feature importances between the `sklearn` and **R** implementations of the random forest when the algorithm is trained on the multiclass realization of the data. Labelled datapoints with ”ff” indicate that a feature is a photon flux fraction (relative to the total photon flux across the *Chandra* band) measured in a particular energy range (in keV). Feature importances are ranked using the mean decrease in the Gini coefficient as described in Section 3.6. Features with a lower mean decrease are less important to the classifier.

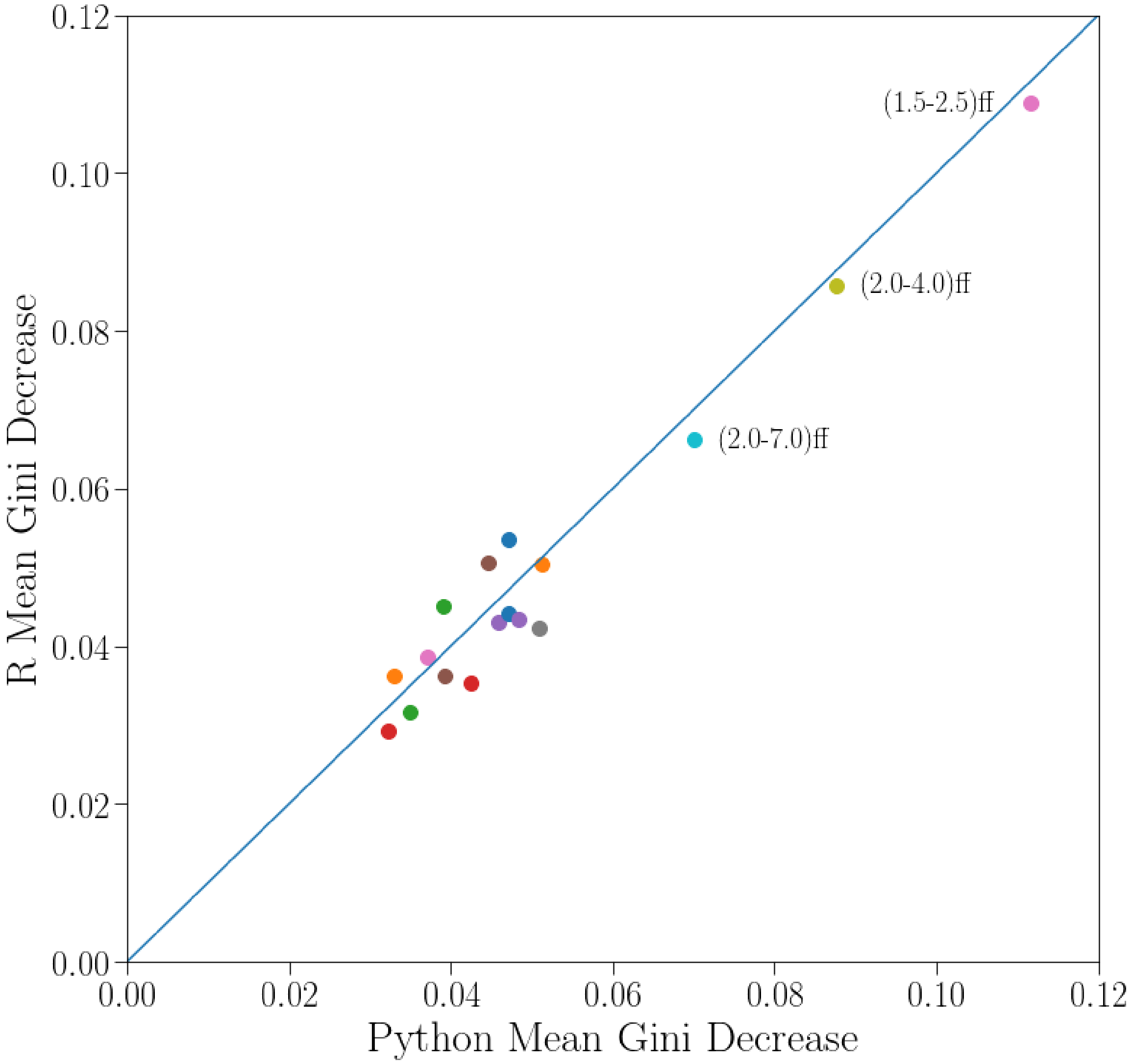

Figure 3.5: Comparison of feature importances between the `sklearn` and **R** implementations of the random forest when the algorithm is trained on the binary realization of the data. Labelled datapoints with "ff" indicate that a feature is a photon flux fraction (relative to the total photon flux across the *Chandra* band) measured in a particular energy range (in keV). Feature importances are ranked using the mean decrease in the Gini coefficient as described in Section 3.6. Features with a lower mean decrease are less important to the classifier.

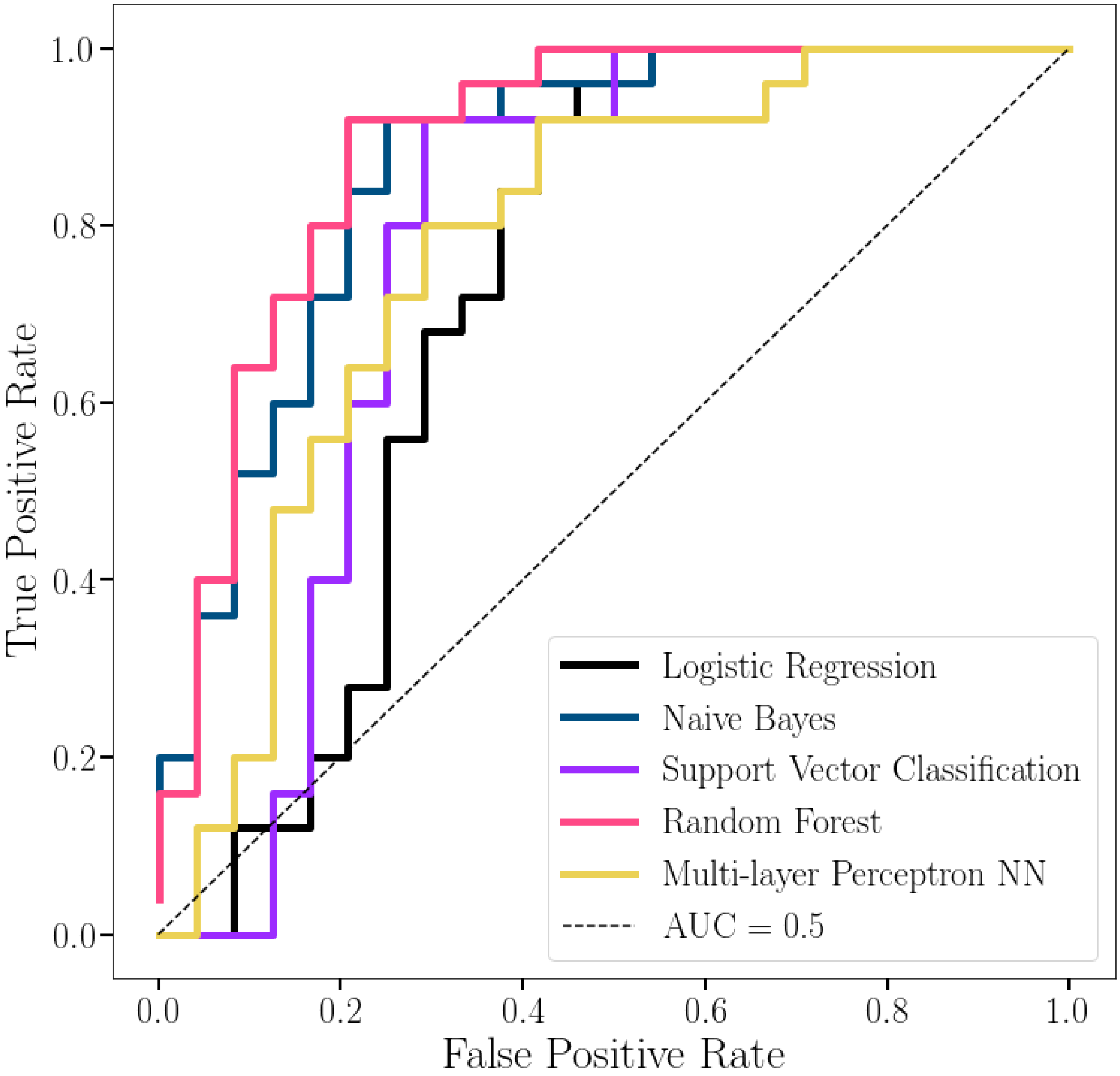

Figure 3.6: Receiver Operating Characteristic (ROC) curve for all sklearn algorithms when trained on the binary realization of the dataset. An uninformative classifier is plotted for reference.

to be insufficient as a discriminating classifier based on theoretical models of emission for different X-ray emitters. However, our results seem to suggest that for the particular case of separating out XRBs, we may be able to use X-ray information to find the best candidate XRBs. Traditional methods using X-ray colour-colour diagrams or colour-colour intensity diagrams have hinted at partial separability between XRBs and other kinds of objects, though often with significant overlap and a dependence on the energy range available (Prestwich et al., 2003). The hard X-ray range of telescopes like *NuSTAR* has been shown to improve separability (Vulic et al., 2018). Our results suggest that using a higher-dimensional approach with this method may yield more useful results. In addition, the use of ML techniques does not require the same kind of modelling assumptions as ordinary regression in determining, for example, a linear decision boundary in a colour-colour plot.

Across all of our modes of analysis and performance evaluation (accuracy, recall, CV score, AUC, and confusion matrices), the random forest algorithms (both sklearn and R) tend to give overall the best performance. The superior performance of the random forest is unsurprising, given the following properties of our dataset:

- Few classified examples (< 200 total, some categories with fewer than 20 members)
- Relatively few features (< 20, a number of which are linearly dependent on each other)
- Features that have differing normalizations (most are ratios but there are 3 features that are bounded differently)
- Complex feature space that is unlikely to be linearly separable by category (significant overlap in the feature space between different kinds of object)

Logistic regression tends to work well when feature values make independent, additive contributions to class probabilities. Naive Bayes makes the assumption that features are conditionally independent from each other. Multi-layer perceptrons often require significant tuning of hyperparameters in order to be useful (Pedregosa et al., 2011). Support vector classifiers do not have these same limitations, however they do not provide direct computation of class membership probabilities (in `sklearn` these are computed from five-fold cross validation instead). The performance of our SVC could likely be improved through tuning of algorithm hyperparameters, including those which account for class imbalance.

Using our best performing algorithm, namely the `sklearn` random forest, the prediction of the 780 unclassified sources using the multiclass approach results in 345 candidate XRBs, 321 candidate background AGNs, 101 candidate foreground stars, and 13 candidate supernova remnants. However, not all of the candidate XRBs are equally likely - the probability inferred from random forest is based on the fraction of decision trees that vote for a particular classification. Amongst the candidate XRBs identified by the multiclass random forest approach, 19 have a probability of 90% or greater.

Conversely, the binary approach identifies only 217 candidate XRBs, while the remaining 563 sources are classified as non-XRB. Of the candidate XRBs, 16 have a probability of XRB classification (as decided by the random forest) of 90% or greater, all of which are included in the set of 19 high-probability candidates from the multiclass approach. The difference in the number of high-probability XRB candidates can mostly be explained by the difference in classification thresholds. By default, in the binary approach, a source requires 50% + 1 trees to classify it as an XRB to meet the threshold for XRB classification. By contrast, in the multiclass approach, a source requires only a plurality of trees to classify as an XRB, which can be a smaller fraction of trees overall. The difference in the number of high-probability classifications for the binary versus multiclass approach is due to the binary approach giving more accurate classifications overall.

One advantage of the random forest method is that it is straightforward to obtain the relative importance of different features to the final, trained classification algorithm using methods such as the mean decrease in Gini coefficient. The Gini coefficient measures the homogeneity of a group of objects. The mean decrease of the Gini coefficient is a measure of the decrease in the homogeneity of objects in each node of the decision tree when a particular feature is removed from all of the trees in the forest (Breiman et al., 1984). Sklearn and **R** both provide this measurement of the importance of features in random forest models, and we have plotted this value (normalized to sum to 1 for all features) for the sklearn and **R** random forests in both the binary and multiclass approach in Figures 3.4 and 3.5.

We find that although there is variation between the two algorithms, both sklearn and **R** tend to weight the same features roughly as important, and in the binary approach (which we expect to be less noisy overall), we find even stronger agreement. The parameter space for a few of the most important features in the multiclass and binary cases is shown in Figures 3.2 and 3.3. As expected, even when plotting the most significant features, there is no clear separating boundary between the categories of sources. In the multiclass plots, we note that supernova remnants and foreground stars tend to have large fractions in the softest bands (those that begin at the 0.5 keV edge of the *Chandra* range), while AGN are found throughout the parameter space. XRBs and XRB candidates also occur throughout the parameter space, however they tend to cluster at intermediate values of these soft bands. In the binary case, the strip in the parameter spaces occupied by XRBs is narrower than in the multiclass approach. The binary approach-classified XRBs and XRB candidates tend to have some flux in the harder bands (e.g., 2.0–7.0 keV), but they tend neither to be the softest nor the hardest sources in the sample.

Curiously, we find that the features judged to be strongly predictive are flux ratios obtained from bands that are less common in traditional hardness ratio analyses, and are not generally measured for *Chandra* datasets except in the ACIS EXTRACT defaults. Some of these bands, such as 1.7–2.8 keV or 0.5–1.0 keV, tend to be narrower than typical cuts made for flux ratios, whereas others, such as 2.0–4.0 keV and 2.0–7.0 keV, are atypical cuts made for flux ratios even if they are relatively wide slices of the *Chandra* energy range. Narrower bands are expected to be less useful overall since more counts are needed in order to measure the flux in these bands accurately. Detailed interpretation of the significance of these bands is deferred to future work, though we briefly place the bands selected in context here.

Prestwich et al. (2003) plotted different categories of X-ray source in nearby galaxies (including M31's bulge) observed with *Chandra* using hardness ratios with separations at 0.3–1.0 keV, 1.0–2.0 keV, and 2.0–10.0 keV. Several properties of X-ray sources are noted. LMXBs tend to have spectra described by a power law of photon index 1.5–2 combined with intrinsic absorption. HMXBs appear harder in the *Chandra* range, with an index of 1–2, though there is a dependence on the accretor type; neutron star XRBs typically have harder spectra than black hole XRBs due to neutron stars having a solid surface (Binder et al., 2015). Both classes can have emission above 10 keV, though LMXBs and HMXBs are expected to peak above and below 15 keV.

Combined, both classes of XRB are intrinsically harder than supernova remnants, which have soft spectra peaking below 2 keV because of a combination of shock-heated plasma and atomic emisison lines (Yokogawa et al., 2003). X-ray active foreground stars, likewise, also have spectra dominated by optically thin thermal plasma and are expected to show up as relatively soft X-ray emitters (Ducci et al., 2013). It is likely that low values for the 2.0–7.0 keV and 2.0–4.0 keV bands and high values for the 0.5–1.0 keV or 1.7–2.8 keV bands are being driven by the intrinsically soft SNR or foreground star X-ray spectra. An interesting future study would be to combine SNRs and foreground stars into a single category and use a three-class approach to training and classification. Binder et al. (2015) used this approach on the basis of their similarly expected properties to classify X-ray sources in NGC 55, NGC 2403, and NGC 4214. Notably, their Bayesian approach also included source position relative to the galaxy, which was not available to us since our source coverage is limited to a handful of fields within M31.

The comparison of AGN to XRBs presents a more complicated picture. AGN are described by a power law continuum with index *Gamma* = 1.7–2 in soft X-rays, with additional contributions from non-thermal inverse Compton scattering of accretion disk photons by hot electrons above the disk, photoelectric absorption edges due to gas along the line of sight, and relativistically broadened emission lines around 6 keV (Nandra & Pounds, 1994; George & Fabian, 1991; Morrison & McCammon, 1983; Fabian et al., 2000). It is possible that the relative importance of the 2.0–7.0 keV band in our AGN/XRB discriminiation is due to the relativistically broad iron line present in AGN.

3.7 Conclusions and Future Work

- We have constructed a proof of concept method for improving classification of X-ray sources in nearby galaxies using machine learning techniques.
- After testing a variety of algorithms, we find that random forest classification tends to outperform other classifiers, offering an accuracy of $\sim 85\%$ at separating X-ray binaries from other kinds of contaminating X-ray sources.
- Using a binary approach to classification, we find 16 new strong XRB candidates that are suitable as candidates for followup.
- Cross-matching previously unclassified sources with sources classified using the PHAT survey, we find 5 sources associated with PHAT-identified clusters, of which 4 are classified as XRBs by our random forest algorithm.
- We have found that the X-ray bands which tend to rank highest in importance for classification are typically narrower and/or less commonly used bands, such as the 1.7–2.8 keV, 0.5–1.0, 2.0–4.0, and 2.0–7.0 keV photon flux ratios.

A primarily limitation of machine learning techniques is that they tend to offer poor predictive performance for small sample sizes. In our sample, we have fewer than 200 total classified examples, and there is significant imbalance between the four types of X-ray source identified by previous surveys. The number of sources required for accurate classification depends on the desired significance threshold, amongst other parameters (Beleites et al., 2012). Additionally, classified samples to use as input for ML methods in astronomical data are typically those objects which are brightest, nearest, and have the longest duty cycles, which may impact new, unclassified samples, if they are distributed in areas of the parameter space where there are few examples available for classification (Richards et al., 2006).

The most obvious pathway for improvement of our methodology is to include more classified examples in the algorithm training. For M31 specifically, a recent *Chandra* survey of M31 overlaps with PHAT with 170 new detected sources, which could provide additional classified examples through matching with PHAT-based classifications (Williams et al., 2018). More broadly, a more useful version of this algorithm would be able to predict X-ray classification for any X-ray source, regardless of expected population. Promisingly, our best performance RF algorithms do not strongly weight the only distance-dependent feature in our feature vector (0.5–8.0 keV photon flux) as a strong determinant in classifying X-ray sources. As such, combining results from surveys of well-studied nearby galaxies will improve our detection algorithm, though caution must be taken. For example, the effect of Galactic extinction along different lines of sight may necessitate additional corrections for absorption so that X-ray sources from different galaxies can be compared to each other. Other effects which we could take into account for future trained versions of this algorithm would include treatment of uncertainty on the flux ratios used as features, corrections for variability, and consideration of other features such as spatial location relative to galaxy structure and the properties of optical counterparts (e.g, magnitudes in the various PHAT filters for the M31 dataset) as features. When compared with XRBs and AGNs, neither foreground stars or SNRs show rapid variability (Binder et al., 2015), so a feature which characterizes variability over all observations may be an additional discriminant. Large catalogues of X-ray sources, such as the *Chandra* Source Catalog (Evans et al., 2010), may also provide a mineable source of classified examples.

We could also improve our classification strength by performing a more detailed investigation of algorithm performance for different values of the algorithm hyperparameters. A deeper study would involve conducting a gridded search that varies all algorithm hyperparameters at once and evaluates algorithm performance at each value using cross-validation or a similar performance metric. Although each particular algorithm can be fit to a particular training set quickly, the number of fits required to do a grid search scales as $\prod_{i=1}^{n} f(i)$, where i is the number of hyperparameters and $f(i)$ is the number of values tested for the i^{th} hyperparameter. This can easily become computationally time-consuming to evaluate for all of the algorithms presented in this work. However, a potential future extension would be to perform a more detailed study of the hyperparameter space of the random forest algorithm, since it offers the best performance in this classification task. As discussed above in Section 3.3, the main source of poor classification accuracy in the multi-class approach is misclassifications of AGN combined with low numbers of foreground stars and supernova remnants. A next generation of ML algorithm which includes multiwavelength properties would hopefully rectify this issue – foreground stars, for example, can often be ruled out on the basis of their extreme optical/X-ray flux ratios. Improved classification and identification of these unknown X-ray sources will enable better understanding of the populations of XRBs inside of galaxies, and may also provide clues about the nature of XRB emission.

Bibliography

Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C., & Popp, J. 2012, ArXiv e-prints. `https://arxiv.org/abs/1211.1323`

Binder, B., Williams, B. F., Eracleous, M., et al. 2015, AJ, 150, 94, doi: `10.1088/0004-6256/150/3/94`

Bodaghee, A., Tomsick, J. A., Rodriguez, J., & James, J. B. 2012, ApJ, 744, 108, doi: `10.1088/0004-637X/744/2/108`

Brehm, D., Kargaltsev, O., Rangelov, B., Volkov, I., & Pavlov, G. G. 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts #223, 255.13

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. 1984, Classification and regression trees (Wadsworth & Brooks/Cole Advanced Books & Software)

Casares, J., Jonker, P. G., & Israelian, G. 2017, X-Ray Binaries, ed. A. W. Alsabti & P. Murdin, 1499

Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012, ApJS, 200, 18, doi: `10.1088/0067-0049/200/2/18`

Dorn-Wallenstein, T., Levesque, E. M., & Ruan, J. J. 2017, ApJ, 850, 86, doi: `10.3847/1538-4357/aa9329`

Ducci, L., Sasaki, M., Haberl, F., & Pietsch, W. 2013, A&A, 553, A7, doi: `10.1051/0004-6361/201321035`

Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, ApJS, 189, 37, doi: `10.1088/0067-0049/189/1/37`

Fabian, A. C., Iwasawa, K., Reynolds, C. S., & Young, A. J. 2000, PASP, 112, 1145, doi: `10.1086/316610`

George, I. M., & Fabian, A. C. 1991, MNRAS, 249, 352, doi: `10.1093/mnras/249.2.352`

Gilfanov, M. 2004, Monthly Notices of the Royal Astronomical Society, 349, 146, doi: `10.1111/j.1365-2966.2004.07473.x`

Gopalan, G., Vrtilek, S. D., & Bornn, L. 2015, ApJ, 809, 40, doi: `10.1088/0004-637X/809/1/40`

Gordon, K. D., Bailin, J., Engelbracht, C. W., et al. 2006, ApJ, 638, L87, doi: `10.1086/501046`

Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2003, MNRAS, 339, 793, doi: `10.1046/j.1365-8711.2003.06224.x`

Guillot, S., Rutledge, R. E., Bildsten, L., et al. 2009, MNRAS, 392, 665, doi: `10.1111/j.1365-2966.2008.14076.x`

Huo, Z.-Y., Liu, X.-W., Xiang, M.-S., et al. 2015, Research in Astronomy and Astrophysics, 15, 1438, doi: `10.1088/1674-4527/15/8/023`

Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2015, ApJ, 802, 127, doi: `10.1088/0004-637X/802/2/127`

Ksoll, V. F., Gouliermis, D. A., Klessen, R. S., et al. 2018, MNRAS, 479, 2389, doi: `10.1093/mnras/sty1317`

Liaw, A., & Wiener, M. 2002, R News, 2, 18

Maccarone, T. J., Kundu, A., & Zepf, S. E. 2003, The Astrophysical Journal, 586, 814

Meusinger, H., Henze, M., Birkle, K., et al. 2010, A&A, 512, A1, doi: `10.1051/0004-6361/200913526`

Morrison, R., & McCammon, D. 1983, ApJ, 270, 119, doi: `10.1086/161102`

Nandra, K., & Pounds, K. A. 1994, MNRAS, 268, 405, doi: `10.1093/mnras/268.2.405`

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Prestwich, A. H., Irwin, J. A., Kilgard, R. E., et al. 2003, ApJ, 595, 719, doi: `10.1086/377366`

R Core Team. 2018, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria

Raudys, S. J., & Jain, A. K. 1991, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13, 252, doi: `10.1109/34.75512`

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, ApJS, 166, 470, doi: `10.1086/506525`

Saeedi, S., Sasaki, M., & Ducci, L. 2016, A&A, 586, A64, doi: `10.1051/0004-6361/201526233`

Sonbas, E., Rangelov, B., Kargaltsev, O., Dhuga, K. S., & Hare, J. 2015. `https://arxiv.org/abs/1505.00216`

Spackman, K. A. 1989, in Proceedings of the Sixth International Workshop on Machine Learning (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.), 160–163. `http://dl.acm.org/citation.cfm?id=102118.102172`

Stiele, H., Pietsch, W., Haberl, F., et al. 2011, A&A, 534, A55, doi: `10.1051/0004-6361/201015270`

Verbunt, F., & Lewin, W. H. G. 2006, Globular cluster X-ray sources, ed. W. H. G. Lewin & M. van der Klis, 341–379

Vilhu, O., & Walter, F. M. 1987, ApJ, 321, 958, doi: `10.1086/165689`

Vrtilek, S. D., & Boroson, B. S. 2013, MNRAS, 428, 3693, doi: `10.1093/mnras/sts312`

Vulic, N., Gallagher, S. C., & Barmby, P. 2016, MNRAS, 461, 3443, doi: `10.1093/mnras/stw1523`

Vulic, N., Hornschemeier, A. E., Wik, D. R., et al. 2018, ApJ, 864, 150, doi: `10.3847/1538-4357/aad500`

Wik, D. R., Lehmer, B. D., Hornschemeier, A. E., et al. 2014, ApJ, 797, 79, doi: `10.1088/0004-637X/797/2/79`

Williams, B. F., Lang, D., Dalcanton, J. J., et al. 2014, ApJS, 215, 9, doi: `10.1088/0067-0049/215/1/9`

Williams, B. F., Lazzarini, M., Plucinsky, P., et al. 2018, ArXiv e-prints. `https://arxiv.org/abs/1808.10487`

Yokogawa, J., Imanishi, K., Tsujimoto, M., Koyama, K., & Nishiuchi, M. 2003, PASJ, 55, 161, doi: `10.1093/pasj/55.1.161`

Yukita, M., Hornschemeier, A. E., Lehmer, B. D., et al. 2016, ApJ, 824, 107, doi: `10.3847/0004-637X/824/2/107`

Chapter 4

Probing Galactic X-ray Binaries and Galactic Structure with Gaia DR2

4.1 Introduction

4.1.1 X-ray Binaries

X-ray binaries (XRBs) are rare systems comprised of a main-sequence star in a close binary orbit with a neutron star (NS) or black hole (BH). The accretion of material from the main-sequence companion onto the compact object results in X-ray emission which dominates much of the point source population of the X-ray sky. Aside from the type of accretor, XRBs are principally categorized based on the mass of the companion. Binaries where the compact object accretes from the wind of a star > 10 $M_\odot$ are classified as high-mass X-ray binaries (HMXBs), while those that accrete from the Roche lobe overflow of a < 1 $M_\odot$ companion are known as low-mass X-ray binaries (van Paradijs, 1998; Casares et al., 2017). There are a handful of XRBs where the companion is of intermediate mass $1 - 3$ $M_\odot$, but they are rare compared to the other two types of system. It is expected that many primordial IMXBs have evolved to LMXBs in the present day through mass transfer (Podsiadlowski & Rappaport, 2000).

XRBs are interesting extraterrestrial laboratories that permit the testing of our understanding of physical processes under extremes of gravity, rotation rate, pressure, temperature, and magnetic field strength. In addition, a number of interesting astrophysical phenomena can be studied through XRBs, such as wind physics, neutron star equation of state, and high-energy radiative processes. Aside from their value to these astrophysical questions, XRBs can also provide independent constraints on their formation environment on larger scales (Lehmer et al., 2010; Boroson et al., 2011; Zhang et al., 2012; Tremmel et al., 2013). LMXBs can act as independent tracers of stellar mass, since low-mass stars comprise the bulk of the stellar mass in a population (Gilfanov, 2004). Additionally, LMXBs are preferentially found in areas of high stellar density, such as the globular clusters of the Galaxy and in the direction of the Galactic Center, likely due to their formation by dynamical mechanisms (Clark, 1975; Pooley et al., 2003; Muno et al., 2005; Verbunt & Lewin, 2006; Degenaar et al., 2012). By contrast, the high-mass companions of HMXBs are short-lived, so they are useful for tracing recent star formation in a long-term Galactic evolution context. Observations of nearby galaxies have suggested that the star formation rate of a galaxy scales with both the number of HMXBs and their collective X-ray luminosity, albeit with a moderate dispersion (Grimm et al., 2003; Mineo et al., 2012). Finally, XRBs are one of the few ways to observe the high mass end of the initial mass function (IMF) in an evolved population, since isolated neutron stars and black holes are challenging to observe and study (Verbunt & Hut, 1987; Verbunt, 2003; Dabringhausen et al., 2012).

Although field Milky Way XRBs can often be easier to study because of their close proximity (compared to XRBs in globular clusters or other galaxies), investigating the relationship between XRBs and galaxy parameters for the Milky Way is complicated. Our location within the disk of the Milky Way means that lines of sight where XRBs are expected to be more abundant tend to be heavily extincted.

XRBs tend to have a spatial distribution that is distinct from ordinary stars belonging to the same parent stellar population because the supernova that forms the compact object in an XRB system can impart a velocity kick to the system, often known as a "natal" kick. This velocity kick has two effects: it gives the XRB system a peculiar velocity relative to galactic rotation, and it can substantially displace the system (depending on XRB type) from the star forming region where its progenitor formed (González Hernández et al., 2005; Dhawan et al., 2007). Repetto et al. (2012) investigated how natal kicks at the birth of black hole LMXBs are necessary to explain their observed distribution in the Milky Way, particularly the presence of LMXBs at significant (1 kpc) distances above the disk. They found that these kicks tend to be similar to those found for neutron stars, a property which has been interpreted as a consequence of the asymmetry of the supernova explosion (Janka, 2013).

Naively, we expect that if HMXBs are correlated with star formation on a global scale, they should have a spatial correlation with the sites of star formation in the spiral arms. The shape and extent of the Milky Way's spiral arms is not easy to resolve compared to external galaxies observed face-on. Positions of the spiral arms themselves are typically inferred through the fitting of analytical models to an ensemble of observational tracers, including CO maps, HII regions, pulsars, masers, stellar kinematics, and dust emission (Vallée, 2014). To date, investigations of the correlation between HMXBs and the spiral arms have been done using only two proxies of the spiral arms. Bodaghee et al. (2012) measured spatial cross-correlation between OB associations and HMXBs, finding that they have a characteristic offset of 0.4 ± 0.2 kpc, which is attributed to natal kicks received by HMXBs at their formation. However, by the same models they find no correlation between either OB associations or HMXBs and the spiral arms themselves, which is unexpected given that OB associations are expected to trace out the spiral arms (Brown et al., 1999). Coleiro & Chaty (2013) investigated the spatial relation between HMXBs and star forming complexes (SFCs) finding that they are correlated on two characteristic scales: 0.3 ± 0.05 kpc and 1.7 ± 0.3 kpc, which they interpret as the cluster size and cluster separation, respectively. They also derive a mean migration distance for HMXBs of roughly 0.1 pc and mean migration ages of around 50 Myr (depending on HMXB type) though they note that sample sizes are small and uncertainties are large. A large source of that uncertainty lies in the difficulty in determining distances to XRBs within the Milky Way.

4.1.2 X-ray Binary Distances

A principal reason for desiring accurate distances to XRBs in the Milky Way is that many of these XRBs that can be studied in detail. With the exception of XRBs located in the direction of the Galactic Center, in the Milky Way the population of XRBs can be studied to fainter X-ray luminosities, and identifications of a unique optical counterpart are more straightforward. Since individual XRBs are most easily studied in the Milky Way, our understanding of individual XRBs in other galaxies and their parameters as an ensemble population are affected by studies of nearby XRBs. Measuring the distance to individual XRBs accurately is important because the uncertainty on a number of desired properties in an XRB system can be limited by the uncertainty on distance. For example, measurements of distance can affect the inferred size of the accretor (i.e., neutron star radius), inferred mass of either component of the system (either the companion mass or the mass of the accreting neutron star/black hole), inferred mass transfer rate, and other relevant accretion physics due to the inferred luminosity (Galloway et al., 2003; Nättilä et al., 2017; Steiner et al., 2018).

Outside of the field of the Milky Way, distances to XRBs are determined by associating the XRB to the population that it is embedded in. For example, distances to XRBs in globular clusters tend to have lower relative uncertainty due to the relatively well-determined distances to those clusters. Globular cluster distances are inferred using a variety of population-based methods, including measuring the period of RR Lyrae stars and applying a period-luminosity relation, horizontal branch star fitting, main sequence fitting, and white dwarf sequence fitting (see Chaboyer 1999 for a review). The preferred method is typically measuring the mean magnitude of the horizontal branch stars in the cluster (Harris, 1996).

In the Milky Way field population, measuring distance with a low uncertainty is considerably more challenging. The principal difficulty in measuring distances to XRBs is that they lack a universal property or characteristic that would allow them to be used as a standard candle. In contrast with an ordinary star or even a binary system of ordinary stars, it is considerably more difficult to model the expected luminosity of an XRB and hence derive a distance. XRBs are also extremely rare compared to ordinary stars, meaning that population-based methods of determining distances to objects, such as main sequence fitting of a star cluster, cannot be used on XRB populations. Although one can use the main sequence of ordinary stars in a cluster to determine the distance to XRBs in that cluster, the rarity of XRBs means that constructing an “XRB main sequence” is untenable. Many XRBs will show an excess in bluer filters in the visible domain due to emission from a hot accretion disk, which may include a continuum excess and the presence of emission lines (for an illustrative example see the case of IGR J17329-2731 as studied by Bozzo et al., 2018). X-ray emission from the accretor which irradiates the companion may modify the expected emission at longer wavelengths (Phillips et al., 1999; Muñoz-Darias et al., 2005; Linares et al., 2018). Failing to account for these effects on the expected optical emission of an XRB may lead to incorrect estimates of distance from photometric methods. These effects are themselves modified by the mass transfer rate, accretion geometry, orbital phase, and accretion state of the system, meaning that they can change with time and may require simultaneous multiwavelength observations for distances to be usefully constrained.

A number of techniques have been used to constrain distance measurements of Milky Way XRBs. The most common of these is to measure a photometric distance by assuming that the emission is dominated by the companion at longer wavelengths. In general, this method is subject to substantial uncertainties, not only due to the contribution of the accretor, but also due to uncertainties in spectral classification and calibrating the absolute magnitude (Reig & Fabregat, 2015). A small number of XRBs have had their distances determined via radio parallax or the proper motion of a launched jet (Hjellming & Johnston, 1981; Bradshaw et al., 1999; Miller-Jones et al., 2009). This form of measurement provides relatively tight constraints on distance, but is possible only for objects that are sufficiently radio-bright and moderately nearby.

An X-ray specific method of measuring distances is to use the observed flux from Type I X-ray bursts. These bursts occur when a sufficient amount of accreted material, mostly hydrogen, accumulates on the surface of a neutron star to trigger a thermonuclear runaway that produces a characteristic burst (Lewin et al., 1993). The burst is specifically the result of nuclear burning on the neutron star. Since black holes lack a surface, only neutron stars exhibit these bursts. A subset of these bursts have steady hydrogen burning followed by ignition of a helium layer beneath the hydrogen layer on the surface. The ignition of this helium layer produces a burst that is sufficient to lift the photosphere off the surface of the neutron star. These bursts are known as photospheric radius expansion (PRE) bursts, and the luminosity of the X-ray burst is expected to be at the Eddington luminosity during the expansion and contraction of the photosphere (Kuulkers et al., 2003). Since the Eddington limit is fixed for a particular accretor mass (and gas composition/opacity), this means that the mass, radius, and distance of a neutron star can be constrained by comparing the observed flux to the modelled Eddington luminosity for that object. (Strohmayer & Bildsten, 2006; Bhattacharyya, 2010). The use of X-ray bursts to infer distance was suggested not long after the detection of such bursts by early X-ray satellites, and historically this relation has been calibrated using X-ray bursts observed in Galactic globular clusters (van Paradijs, 1978, 1981; Verbunt et al., 1984). Several Galactic XRBs that exhibit either PRE or PRE-like bursts have had their distances constrained using this method (Basinska et al., 1984; Galloway et al., 2003; Jonker et al., 2004). Evaluations of this method have shown that uncertainties around the modelling assumptions in this method can result in uncertainties in distance, neutron star mass, and neutron star radius (Galloway et al., 2008b).

With the exception of Type I X-ray bursts, most of the distance-determination techniques require the identification of an optical/infrared counterpart to the X-ray source. Identification of a counterpart requires high spatial resolution and accurate determination of X-ray position. However, the large angular size of the Milky Way would require a survey with large coverage for an comprehensive census of Milky Way XRBs. To date, most all-sky or large-coverage X-ray surveys have low spatial resolution (typically ~ arcminutes). The low resolution of such surveys will not, in general, yield a unique optical counterpart for each object. Many X-ray binaries exhibit variability and may not be detectable when re-observed. Existing catalogues of XRBs include sources which have not been re-detected since their discovery prior to the era of high angular resolution telescopes, and as such have poorly-determined positions that could have many candidate counterparts. The presence of interstellar extinction along particular lines of sight can interfere with the identification of optical counterparts for many XRB sources. Aside from studies of individual objects using telescopes such as the *Hubble Space Telescope*, the principal existing parallax survey of objects in the Milky Way was conducted by the *Hipparcos* satellite (Perryman et al., 1997). *Hipparcos* provides parallax for only ~ 100000 sources, and has a fairly shallow limiting magnitude of 12. A handful of nearby XRBs have had their distances determined via *Hipparcos* parallax (see, for example, Chevalier & Ilovaisky 1998). *Hipparcos* data provides reliable measurements of distance within a few hundred parsecs of the Sun, which excludes (based on estimates using the other distance methods described above) the overwhelming majority of XRBs known in the Milky Way.

4.1.3 Gaia DR2 as a Probe of XRB Distances

The successor to *Hipparcos*, the Gaia satellite, was launched in 2013 and aims to have full 5-parameter measurements (position, proper motion, parallaxes) for $\sim$ 1 billion stars and parallaxes accurate to 10% for approximately 100 million sources by the end of its five-year mission (Gaia Collaboration et al., 2016a). To date, there have been two major data releases of Gaia results, the most recent of which was released in April 2018 and is based on the first 22 months of data taken (Gaia Collaboration et al., 2016b, 2018). The current data release (DR2) contains over 1.3 billion sources which have full 5-parameter measurements, an improvement of 5 orders of magnitude of *Hipparcos* for parallax measurements. Depending on the required uncertainties, Gaia DR2 contains measurements down to a limiting G magnitude of 17–21. So far, Gaia DR2 has already provided a wealth of information for studying populations in and nearby the Milky Way that deviate from the expected dynamics of ordinary stars in the Milky Way. For example, measurements of candidate hypervelocity stars using Gaia DR2 have shown that many of them are in fact bound to the Milky Way, but at least one object has an origin in the direction of the Magellanic Clouds, suggesting the presence of a supermassive black hole in the Large Magellanic Cloud (Boubert et al., 2018; Erkal et al., 2018). Gandhi et al. (2018) have already looked for Gaia DR2 candidate counterparts for Galactic black hole transients, finding that distances from Gaia counterparts generally agree with prior distance estimates. Notably, they find that the black hole BW Cir has a Gaia distance of 0.6 ± 0.2 kpc, making it the closest dynamically-confirmed transient black hole, though they acknowledge that this distance is difficult to reconcile with interpretations of the properties of the donor star.

In this work, we seek to expand this investigation by including not only binaries with black holes/black hole candidates but also neutron star/neutron star candidate binaries and those with no clear identification of accretor type. Given that XRBs are expected to deviate from the Milky Way's stellar distribution in subtle to dramatic ways, Gaia offers a unique chance to create a sample of XRBs whose distances are determined by a uniform method that is both more accurate and has better understood systematics than other XRB distance determinations. It also offers an opportunity to calibrate alternative methods of measuring distance for use in the general case where parallax measurements are not available.

4.2 Sample and Methods

Cross-matching XRBs to Gaia requires input catalog(s) of known XRBs and XRB candidates. To date, the most comprehensive catalogs of XRBs in the Milky Way are the Liu catalogs of high-mass and low-mass XRBs. In general, properties of these XRBs (including positional uncertainties) are compiled using the best/most recent (at the time of catalogue creation) observations of these objects. These catalogs are assembled from published observations taken with a variety of X-ray telescopes, including *Uhuru*, *Einstein*, *ROSAT*, *RXTE*, *Chandra*, and *XMM-Newton*. As such, the specific X-ray energies sampled, sensitivities, and coverage of these catalogs is non-uniform. Since the most recent updates to these catalogs were in 2006 and 2007, a number of Galactic XRB candidates discovered since then are missing. However, an advantage of these catalogues is that many of these objects have been studied in detail, especially those with identified counterparts. This implies that the number of expected non-XRB contaminants in this catalogue should be low.

4.2.1 XRB Sample

In order to assemble a sample of XRBs for *Gaia* counterpart matching, we combine the Liu et al. (2006) and Liu et al. (2007) catalogues of Galactic HMXBs and LMXBs. Although the most recent revision of these catalogues is now over a decade old, they still represent the most complete sample in the literature. In total, these catalogues contain 301 XRBs or XRB candidates. The majority of the objects have positional accuracies (equivalent 90 percent confidence) $\sim 1''$or better, typically through an optical counterpart or high-resolution X-ray observation. However, a number of the candidate objects in these catalogues have poorly determined positions, especially those that have not been re-observed since the beginning of the *Chandra* era. We assume that long-wavelength counterparts identified in the catalogues are true counterparts to the LMXB/HMXB or LMXB/HMXB candidates. We have manually removed 2 objects from the Liu catalogs: 1H 0556+286, and 1H 1255-567 (Mu-2 Cru), on the basis that they appear to have been misclassified as HMXBs and are in fact ordinary stars (Berghoefer et al., 1996; Torrejón & Orr, 2001). In order to feasibly attempt to identify *Gaia* counterparts, we select only objects whose positional accuracy is quoted in the catalogues as better than $< 10''$, which provides a sample of 220 XRBs.

4.2.2 Published distance estimates

Distances to XRBs are estimated using many different methods and a goal of this work is to evaluate the quality of these methods. The Liu et al. (2006) and Liu et al. (2007) catalogues provide distance estimates in the notes to the main catalogue files. Although in many cases these are not the only available estimates for these objects, for consistency we adopt them as the "previous distance estimate" and give the original distance reference in our tables. In many cases, only a distance range is quoted in the Liu et al. catalogs and we give the centre of this range. In cases where an upper or lower limit was given, we quote that number as the distance. We note that there are a variety of techniques, instruments, and wavelength regimes used to quantify distance. In addition, uncertainties on distances to XRBs are reported in different ways, including making approximations with no quoted uncertainties. As such, in this work we do not attempt to track the uncertainties associated with previous measurements, except for a handful of cases. In particular, we expect that objects with a radio parallax (VLBI or VLBA) should be more precise than Gaia, and Gaia should agree with parallaxes measured with the *Hipparcos* satellite. However, in general we expect that the distance to Gaia candidate counterparts is more reliable and Gaia's methodology and systematics are better understood than for the ensemble of other methods.

The majority of the objects in our sample have distances measured through the photometry of the companion, using measured apparent magnitude and extinction with an assumed absolute magnitude based on modelling. Many XRBs with a neutron star have had their distance measured using Type I X-ray bursts. Aside from these categories, there are also a handful of objects with *Hipparcos*/radio parallaxes, and a variety of other methods for individual objects. We use the following labels for different distance methods:

- phot: photometric distance using apparent magnitude, extinction, and assumed absolute magnitude of companion
- SEDfit: broad-band SED is fit to an assumed model of the companion star/accretion disk with distance as a fitted parameter
- A_V: distance measured using extinction models/galactic column density
- jetPM: distance measured using jet proper motion
- cluster: distance is assumed to be that of an associated cluster/OB association
- burst: X-ray burst is used as a standard candle to obtain distance
- VLBAPLX/VLBIPLX: parallax measured using radio interferometery
- Kin: distance inferred from the kinematics of associated Hi Regions
- HipPLX: distance measured using parallax from the *Hipparcos* satellite
- unknown: no previous distance measurement

4.2.3 Cross-matching

We searched for counterparts to our XRB sample by cross-matching with the Gaia DR2 public release. Initially, we collected all potential counterparts with a tolerance of $< 10''$and then refined the matches to only include counterparts whose angular separation was less than the quoted positional uncertainty for each individual object. In the case that an object had asymmetric positional uncertainties in right ascension versus declination, we conservatively chose the maximum of these two. With this refinement, 99 XRBs from the Liu catalogues have at least one candidate Gaia counterpart. In total, we find 126 potential counterparts for the Liu XRBs. Most objects have only 1 counterpart, while a handful (those with more poorly determined positional accuracy) return 2 or more potential counterparts.

We further refined our sample of potential XRB counterparts by considering the probability that each Gaia source is aligned with the position of the XRB by chance alone. To estimate probability of our X-ray sources matching a random Gaia source, we picked 5000 random coordinates within 0.1 deg of each X-ray source and measured the distance between these random coordinates and the closest real Gaia source. We approximated the probability of a random match by the fraction of random points which are located within a distance of a Gaia source equal to the separation between the X-ray source and the candidate Gaia counterpart. After removing the counterparts with a probability of chance overlap greater than 10%, we obtain 88 Gaia candidate counterparts to the Liu XRB sample, the majority of which have reported parallaxes. At this level, only two objects have more than one potential Gaia counterpart: AX J1639.0-4642 and SAX J1711.6-3808. Each of these objects has one potential counterpart with a parallax, and one without. For the purposes of simulation/distance comparisons, we use the counterpart that has a parallax (and hence a distance measurement). Twenty of the XRBs with Gaia candidate counterparts have no previous distance measurement (as of the Liu catalogs).

Before proceeding, we consider the potential biases of our sample compared to the unmatched sample. HMXBs are more likely to have a counterpart than LMXBs - there are 187 LMXBs and 114 HMXBs in the Liu catalogs, but we find only 32 Gaia candidate counterparts to LMXBs and 54 Gaia candidate counterparts to HMXBs. In general, objects are more likely to have a Gaia candidate counterpart if they have well-determined positions and/or a bright optical counterpart. Our counterpart matching is also more sensitive to objects that are away from the Galactic Centre and away from the Galactic plane - the fraction of objects in the Liu catalogue that have a Gaia candidate counterpart is higher in directions away from the Galactic Centre.

4.2.4 Distances and Final Sample

To obtain the distance for each counterpart, we match the Gaia source ID to the catalogue of Bailer-Jones et al. (2018), which uses a Bayesian method to infer distances. In this work, we quote distance uncertainties as the 1σ bounds on the posterior probability density function for distance. In general, this function is asymmetric about the peak value, so we have asymmetric error bars. The prior of this Bayesian method models the Galactic stellar density as an exponential disk, so the particular distance prior assumed for each object depends on that object's position in Galactic coordinates. For ordinary stars, information such as line-of-sight extinction, measured T_{eff}, and magnitude/colours in the Gaia filters can provide additional distance constraints. However, for XRBs we prefer the position-plus-parallax-only method used by Bailer-Jones et al. (2018), since modelling the expected value of the additional other parameters in an XRB system is more complex than for an individual star or ordinary binary.

Since LMXBs do not follow the same spatial distribution as the stellar distribution assumed by Bailer-Jones et al. (2018), we must cautiously interpret the distances to LMXBs (Grimm et al., 2002). For example, an exponential disk model would prefer smaller distances for objects along lines of sight that are out of the plane of the Milky Way. However, this may not be optimal for LMXBs, given that they can be displaced from the stellar distribution by supernova kicks.

Of the matched XRBs, 77 of the Liu catalogue counterparts have a parallax. 25 of these counterparts are associated with LMXBs, while 52 are associated with HMXBs. Several of the matched objects have a negative measured parallax. In this case, the distance we obtain is dominated by the assumptions of the prior (see discussion in Luri et al., 2018 and Hogg, 2018). We plot the positions of the Gaia candidate counterparts to the XRBs for face-on and edge-on projections of the Milky Way in Figures 4.1,4.2 and 4.3.

4.3 Results

Several expected results are evident in the edge-on and face-on views of the XRB sample. First, as shown in Figure 4.2, HMXBs appear to trace out the nearby (i.e., within 5–8 kpc) arms of the Galaxy. HMXBs are concentrated primarily along the Galactic disk, while LMXBs can also be found at higher and lower Galactic latitudes. Since HMXB luminosity is correlated with star formation rate in star-forming galaxies, and spiral arms are the primary sites of star formation, it is reasonable to infer that they should be spatially close to spiral arms. Figure 4.1 shows that the LMXBs are correlated with the Galactic Center. This result is also expected since

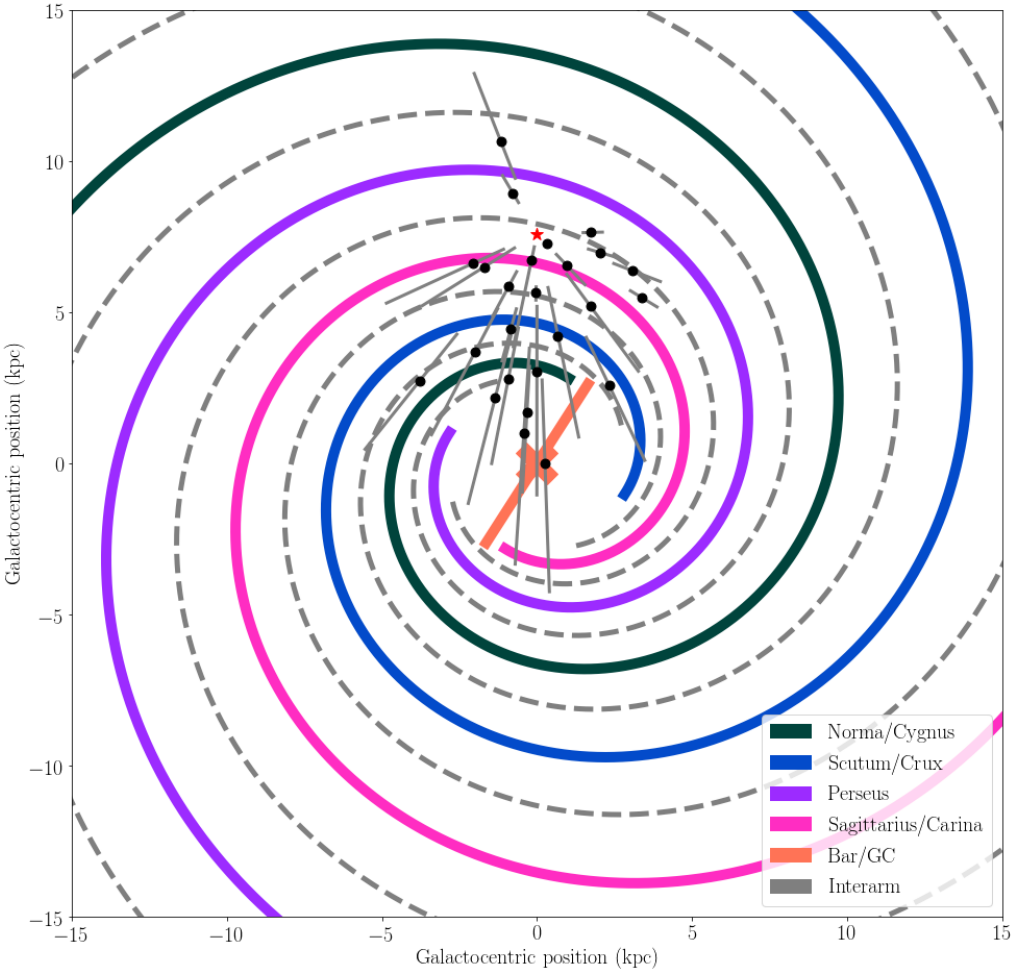

Figure 4.1: Face-on distribution of Gaia counterparts for Liu LMXBs. The spiral arms are modelled using the symmetric spiral arm model of Vallée (2008). Interarm regions are modelled as the symmetric arm model phase shifted by 45 degrees. Error bars for distance/parallax represent the 1σ uncertainties. In this context, the PDF is non-Gaussian so the 1σ region of the probability distribution function is asymmetric about its peak value.

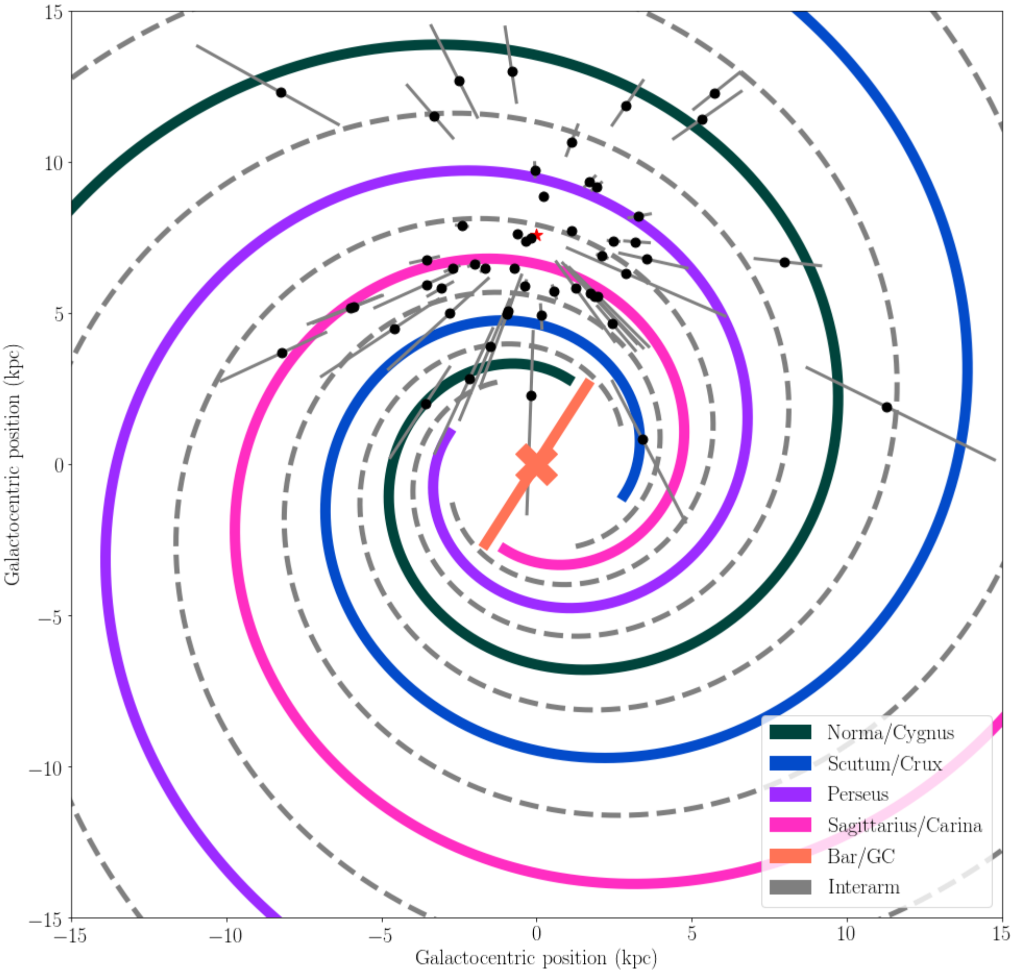

Figure 4.2: Face-on distribution of Gaia counterparts for Liu HMXBs. The spiral arms are modelled using the symmetric spiral arm model of Vallée (2008). Interarm regions are modelled as the symmetric arm model phase shifted by 45 degrees. Error bars for distance/parallax represent the 1σ uncertainties.

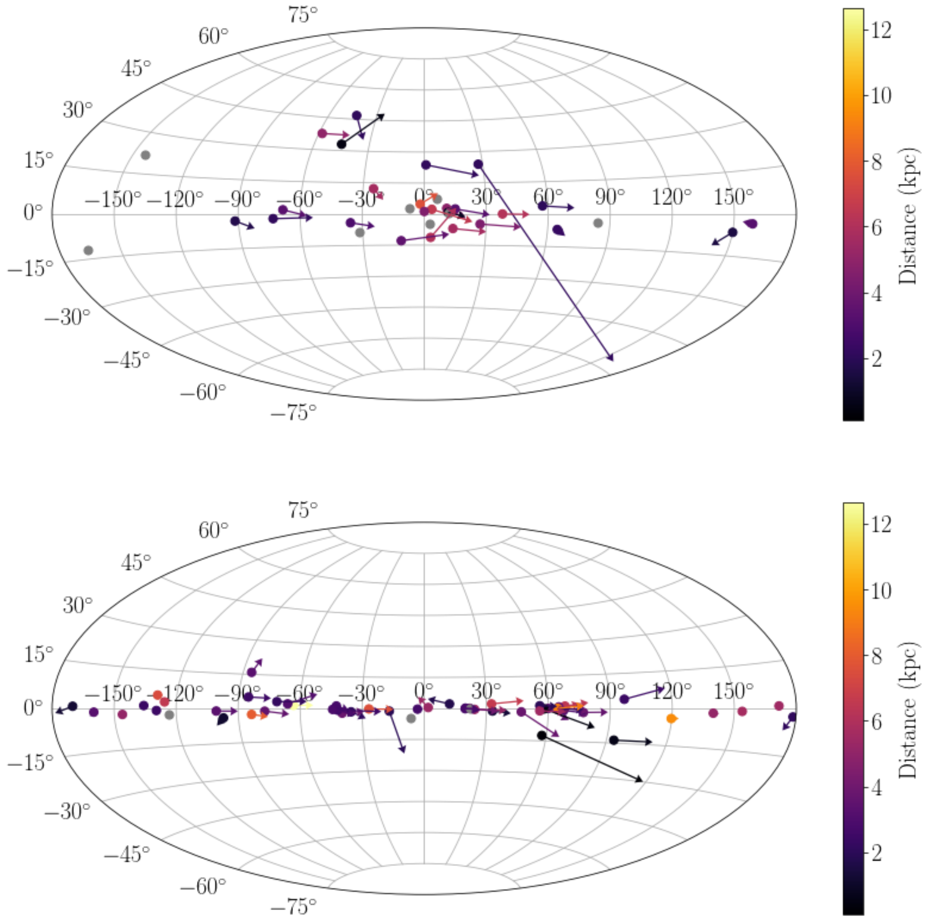

Figure 4.3: Sky position of Gaia candidate counterparts to Liu catalogue LMXBs (top) and HMXBs (bottom) in Galactic coordinates. Vectors represent proper motion of sources as reported in Gaia. The vector length represents relative size of proper motions. Objects which lack a distance measurement via parallax are marked in gray. We omit PM uncertainties here, though we note that the majority of the proper motions have $> 2\sigma$ significance.

we know that LMXBs are tracers of stellar mass and are preferentially formed in dense areas with high stellar encounter rate, such as the Galactic Bulge. Examining the proper motions of our sample as plotted in both panels of Figure 4.3, we see that the proper motions of our candidate counterparts are dominated by Galactic rotation, moreso for the HMXBs than the LMXBs. More of the LMXBs have proper motions not aligned with Galactic rotation. These results are expected, since the 3D motion of XRBs through the galaxy should be governed by a combination of the motion of the stellar population where they formed (i.e., Galactic rotation) and the peculiar velocity added by supernova natal kicks. LMXBs receive, on average, larger natal kicks than HMXBs, which in turn means we should expect their proper motions to be less shaped by Galactic rotation.

4.3.1 Distance Measurement Comparison

Since parallax measurements will not, in general, be available for most Galactic XRBs (we find parallax measurements for less than one third of the combined LMXB/HMXB catalogue), it is useful to use objects with parallax measurement as a diagnostic for other distance methods. We show a comparison of previous distance measurements with those derived from the Gaia candidate counterparts in Figures 4.4 and 4.5. In these plots, we include uncertainties for our Gaia candidate counterparts, but omit the uncertainties on previous measurements given the difficulties in comparing methods, instruments, and the fact that many distances are assumed rather than measured directly. A useful component of this work is to tabulate distance methods for XRBs with Gaia candidate counterparts, since distances in the Liu catalog are reported without specifying the methodology (or whether distances are measured/assumed). Gaia measurements tend to agree strongly with objects whose parallax has been measured either by *Hipparcos* or radio interferometery (VLBI/VLBA). Radio parallaxes are expected to be significantly more accurate than Gaia. Therefore, comparing with radio parallaxes verifies the assumptions of the Gaia prior (at least for the distance ranges and directions where objects with radio parallax are available). In addition, there appears to be no obvious trend in difference between new and previous distances for photometrically-measured distances, aside from objects that are closer being more likely to agree with Gaia distances. However, distances measured using Type I X-ray bursts do show evidence of a trend with a plausible physical interpretation. As shown in Figure 4.4, distances measured using Type I X-ray bursts are systematically larger than those measured via Gaia candidate counterparts. We caution that our uncertainties are still large and our sample size small, but this would seem to suggest that Type I X-ray bursts are intrinsically less luminous than predicted by modelling. This agrees with previous results on systematic biases in distance determination via Type I X-ray bursts. Galloway et al. (2008b) demonstrated that the choice to assume that the touchdown flux (the flux measured when the expanded photosphere of the neutron star touches down back onto its surface) is either at the Eddington flux or sub-Eddington may introduce large systematic uncertainties to distance measurements of X-ray bursting XRBs. Studies of bursting sources using the Rossi X-ray Timing Explorer have indicated that a number of these sources are significantly sub-Eddington in their peak fluxes (e.g., Galloway et al., 2008a).

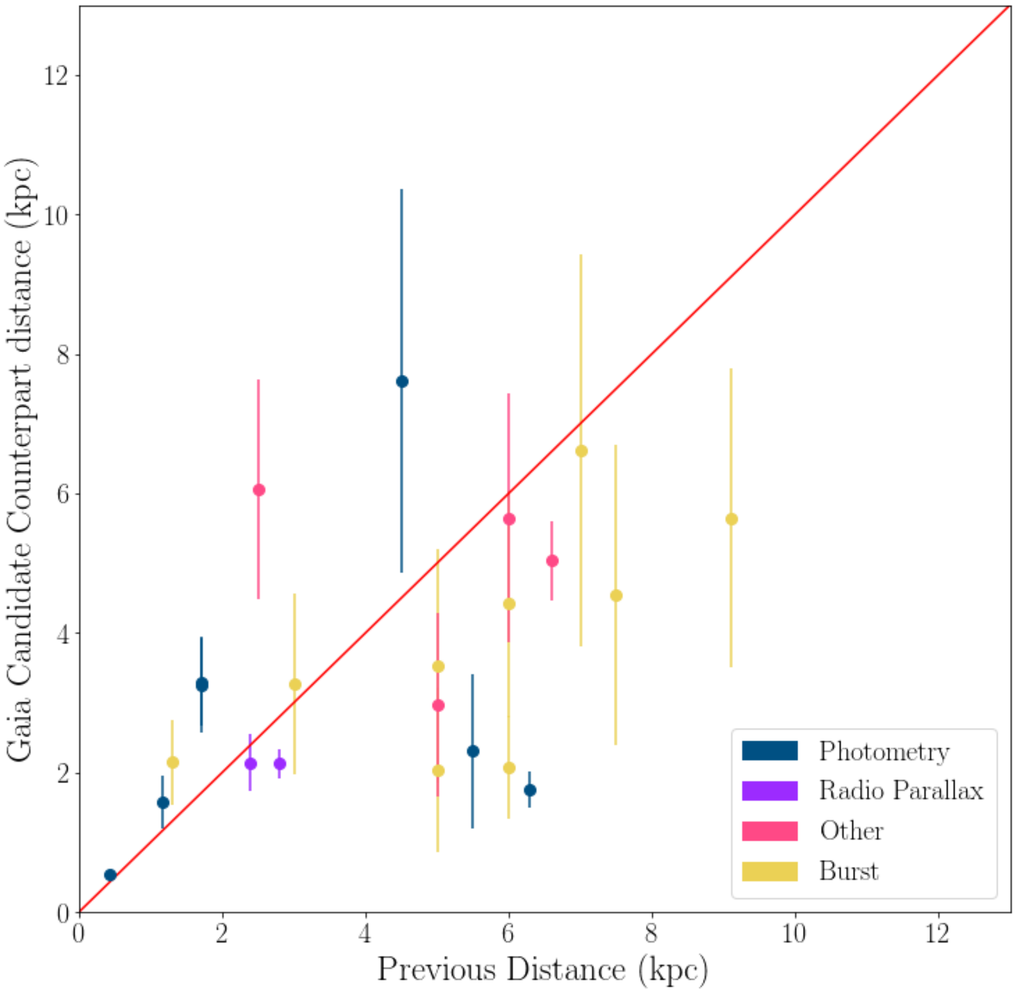

Figure 4.4: Comparison of previous distance measurements against the distances obtained in this work for Liu catalogue LMXBs. Previous distances are obtained from the appropriate literature reference given in the Liu catalogs, while the distance in this work is the distance to the Gaia candidate counterpart for each LMXB.

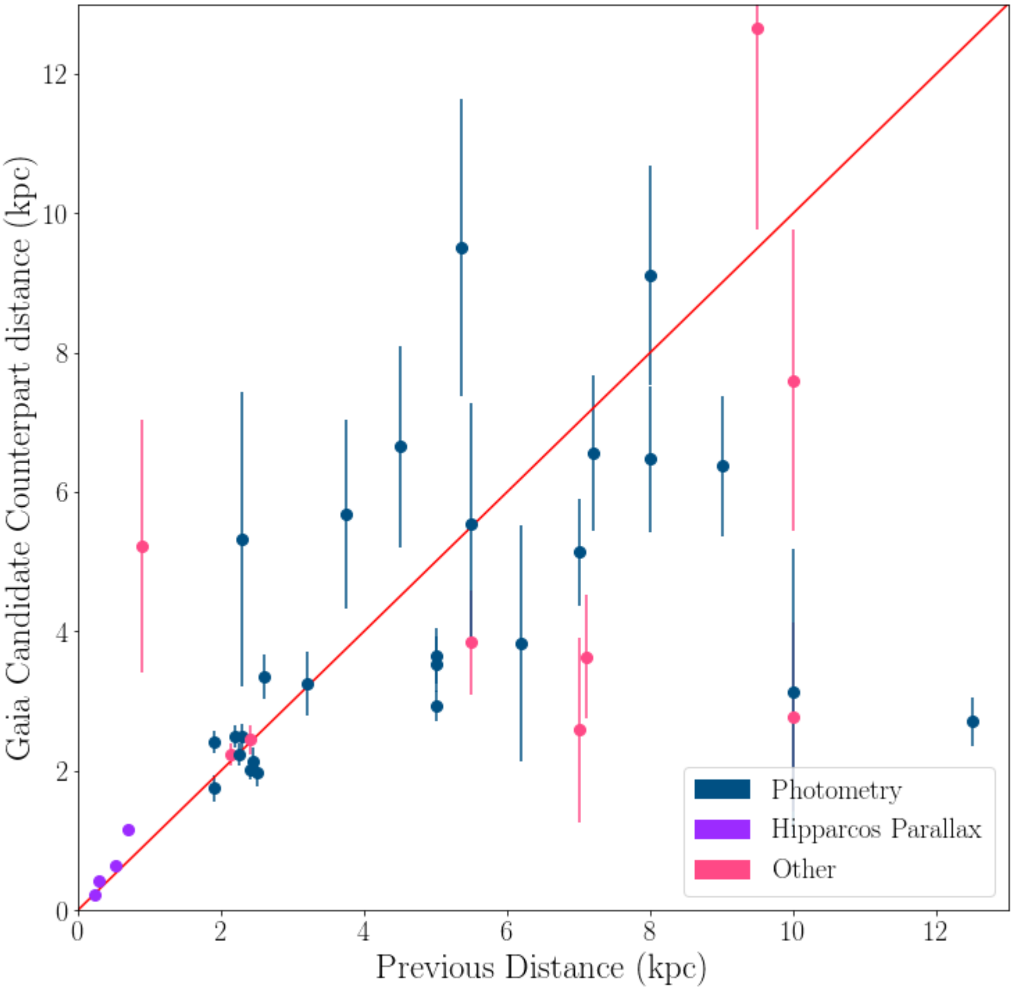

Figure 4.5: Comparison of previous distance measurements against the distances obtained in this work for Liu catalogue LMXBs. Previous distances are obtained from the appropriate literature reference given in the Liu catalogs, while the distance in this work is the distance to the Gaia candidate counterpart for each LMXB.

4.3.2 Spatial Distribution and Spiral Arms

To investigate the relationship between XRBs and galactic structure, we compare the XRB distributions to the spiral arms of the Milky Way. We use the symmetric arm model of Vallée (2008). This model is analytically defined: the precise shape, symmetries, structure, and extent of the spiral arms of the Galaxy are nontrivial to determine due to our location within the Milky Way. This symmetric model is fitted to agree with a variety of observations, including dust, HI GAS, CO gas, and maps of stellar velocities. This model defines the midpoint of 4 identical arms phase shifted by 90°. We further define interarm regions by shifting the existing arms by 45°.

For each XRB, we compute three properties:

1. the two-dimensional distance to the nearest spiral arm for a face-on projection
2. whether the XRB is leading or trailing its closest spiral arm
3. whether the XRB is closer to the midpoint of a spiral arm or the midpoint of an interarm region

Given that many of the uncertainties for the distances quite large, counts of these quantities depend strongly on the posterior distribution function of the distances. In order to assess how much these quantities change, we create 10,000 realizations of the distance for each object using the posterior distribution function defined in Bailer-Jones et al. (2018), and compute the 3 quantities above for each object in each iteration. A sample distribution of distances, for V884 Sco, is given in Figure 4.6.

After computing whether each object is closer to an arm or interarm region, whether it is leading or trailing the nearest spiral arm, and the distance to the nearest spiral arm, we calculate the fraction of objects leading/trailing and fraction of objects close to an arm/interarm for each of the 10,000 runs. Under this construction, since we have effectively partitioned the galaxy into two equally-sized regions (closer to arm/closer to interarm, leading/trailing the nearest spiral arm), we expect the following for the distribution of these fractions: If the distribution of LMXBs/HMXBs fractions peaks at a value greater than 0.5 for a particular structure (arm/interarm/leading edge/trailing edge), then we interpret that LMXBs/HMXBs as being correlated with that structure. Conversely, if the distribution peaks at a value less than 0.5, we interpret LMXBs/HMXBs as being anti-correlated with that structure. If the distribution peaks at 0.5, we interpret LMXBs/HMXBs as being uncorrelated with that structure. We treat the uncorrelated case as the null hypothesis for LMXBs and HMXBs individually.

In each run, we exclude from the fraction any object that lies at a distance of less than 3.1 kpc from the Galactic center, classifying them separately as bulge sources. We choose 3.1 kpc because it is given as the half-length of the bar superimposed on the cartographic plots of Vallée (2008), and it is noted therein that it becomes difficult to separate the beginnings of the spiral arms with the bar itself at approximately this distance. In each run, on average two HMXBs and five LMXBs were classified as bulge sources. The resulting fractions and their uncertainty distributions are plotted in Figures 4.7 and 4.8.

Across the simulation, LMXBs and HMXBs both exhibit a roughly normal distribution in both fractions, though in both the leading/trailing or arm/interarm case, the LMXB distribution possesses a larger spread. To compare these measurements to each other and to the null

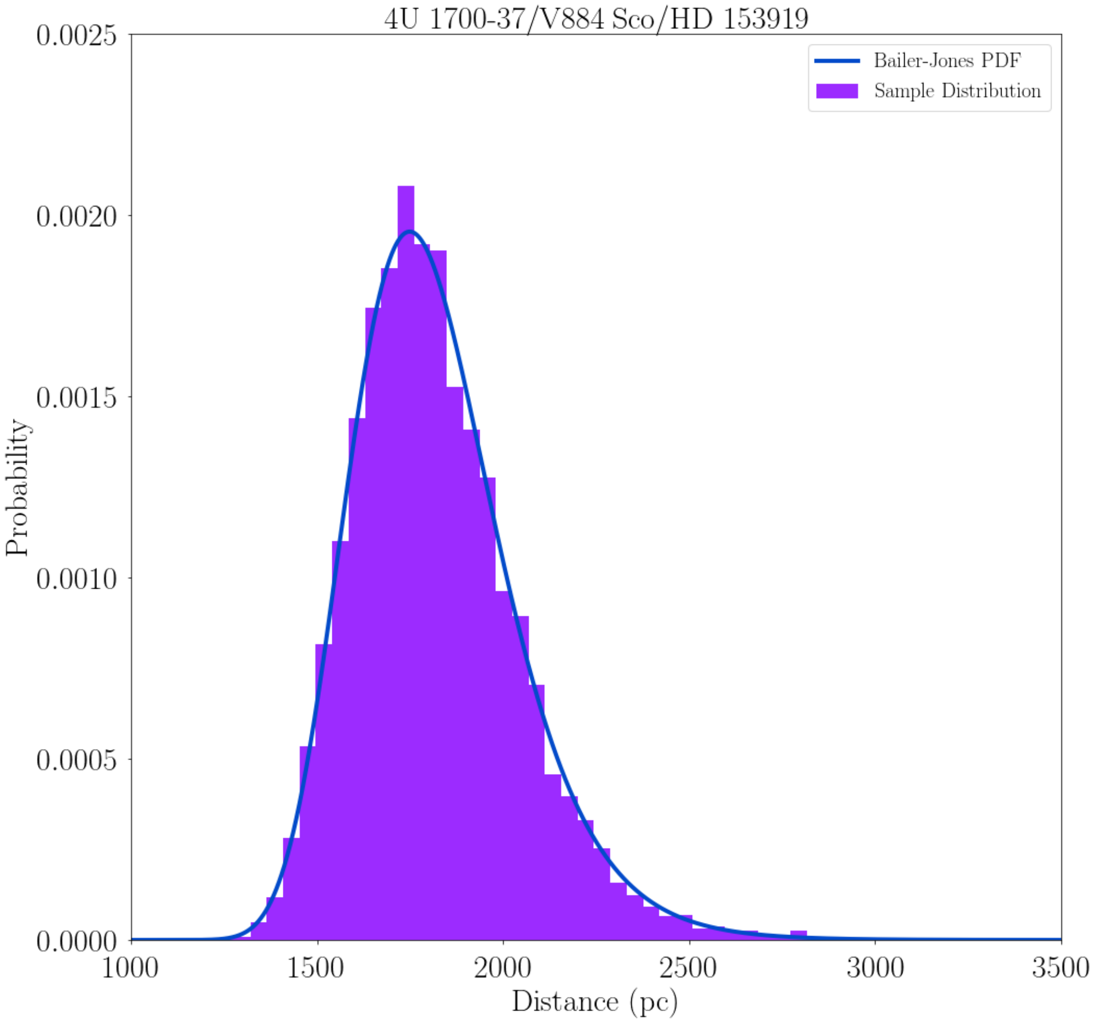

Figure 4.6: Posterior distribution function (PDF) for the distance to V884 Sco using the Bailer-Jones et al. (2018) prior prescription. The posterior distribution function for distance is calculated from the Bailer-Jones prior, the counterpart's parallax, and its parallax uncertainty. The shaded area represents the histogram of distances for 10,000 random draws from this PDF.

hypothesis (that they are uncorrelated with arms/interarms and leading/trailing spiral arms), we tested these uncertainty distributions for normality. Since the interarm/trailing fraction is complementary to the arm/leading fraction, we consider only the arm/leading fractions. The LMXB arm fraction, HMXB arm fraction, and HMXB leading fraction are all considered normal by the D'Agostini K^2 test and the Jarque-Bera test for $p = 0.05$. However, the LMXB leading fraction does not pass either of these tests at $p = 0.05$ Additionally, none of the four distributions are normal according to an Anderson-Darling test at $p = 0.05$. Since the distributions are not truly normal, we report the fraction measurements and their uncertainty in two ways: first using the standard deviation as the 1σ uncertainty, and then reporting the the 95th/5th quantiles as the uncertainty.

Our measurements of the fraction of HMXBs/LMXBs that are correlated with spiral arms/interarm regions yields the following results:

- Fraction of HMXBs that are closer to a spiral arm: 0.54 ± 0.05 (at 1σ), $0.54^{+0.08}_{-0.08}$ at the 95th and 5th quantiles
- Fraction of LMXBs that are closer to a spiral arm: 0.40 ± 0.09 (at 1σ), $0.4^{+0.2}_{-0.2}$ at the 95th and 5th quantiles
- Fraction of HMXBs that are leading the nearest spiral arm: 0.46 ± 0.05 (at 1σ), $0.46^{+0.09}_{-0.09}$ at the 95th and 5th quantiles
- Fraction of LMXBs that are leading the nearest spiral arm: 0.5 ± 0.1 (at 1σ), $0.5^{+0.2}_{-0.1}$ at the 95th and 5th quantiles

We cannot reject the null hypothesis that HMXBs or LMXBs are spatially uncorrelated with spiral arms, at even 1σ, since the errors overlap with $F_{fraction} = 0.5$. We cannot reject the null hypothesis for either HMXBs or LMXBs exhibiting no preference leading or trailing their nearest spiral arm. The LMXB and HMXB fractions also overlap with each other at the 1σ level. LMXBs exhibit a mild preference for being found in interarm regions, while HMXBs show only a mild preference for being found in the spiral arms. LMXBs appear to be uncorrelated with leading or trailing their spiral arm, while at low significance the HMXBs appear to prefer trailing their nearest spiral arm.

In the context of Galactic structure, it is known that HMXBs trace SFR on Galactic scales, so it is reasonable to expect they should trace it on resolved scales in some fashion and should exhibit a distinct spatial correlation (Grimm et al., 2003). Naively it can be assumed that star formation should happen at the leading edge of a spiral arm where the gas accumulates (see Koda et al., 2012 for M51 as an illustrative example of star formation and its relation to spiral arm structure). Taking these assumptions together, HMXBs should be found at the leading edge of spiral arms, and should exhibit a strong preference for spiral arms versus interarm regions. However, we find only a mild preference for spiral arms: the distribution of fractions for the simulation peaks at 54% of the HMXBs being closer to an arm than an interarm region, but the wings of the distribution include the uncorrelated and anti-correlated case. By contrast, we expect that LMXBs should exhibit no strong preference for spiral arms; they represent (collectively) an older population that is also more strongly perturbed by the strength of its SN kicks. Since LMXBs can be much older, it is not expected or required that they are still nearby to the spiral arm that formed their progenitor – there may have been multiple Galactic rotations since the LMXB itself formed. Additionally, their high velocity kicks mean they can be substantially displaced from the star-forming region where they initially formed. This process is already required to explain the presence of LMXBs at high Galactic latitudes where they would not be expected to form *a priori* due to the low stellar density (see, for example, Repetto et al., 2012). Consequently, LMXBs as a population should be uncorrelated with spiral arms since their distribution would be unperturbed by either the presence or absence of spiral arms. This makes our result, which shows LMXBs anti-correlated with spiral arms (though at low significance), difficult to explain.

The lack of strong preference for HMXBs being closely associated with spiral arms could have a number of possible implications:

- Star formation does not occur at the leading edge of spiral arms.
- The time delay between star formation and HMXB accretion starting manifests itself as a spatial separation between the spiral arm and HMXBs due to the pattern speed of spiral arms.
- HMXB natal kicks may be larger than expected.
- Our sample is not large enough and does not have sufficiently small distance uncertainties as an ensemble to measure the correlation we expect from first principles.

Our HMXB sample comprises only ~50 objects, and the uncertainties are still substantial. As such, though we can rule out a very strong spatial correlation or anti-correlation between HMXBs and spiral arms, we cannot use our result to distinguish between the scenarios listed above. Since we are unable to reject the null hypothesis that HMXBs are uncorrelated with spiral arms, our result is consistent with Bodaghee et al. (2012)'s analysis, which found that HMXBs are not spatially correlated with spiral arms. The scale at which the HMXB/SFR correlation breaks down (if at all) is not well-constrained. In nearby galaxies, the X-ray sources are typically studied by considering the integrated properties of the entire population (for example, X-ray luminosity function) and comparing to global parameters of the galaxy. Correlating XRBs with galactic structure is challenging since galaxies that are close enough to resolve on the desired scales require many fields in order to encompass the entire galaxy. In addition, contamination from X-ray sources in front of or behind the galaxy creates additional difficulties. Swartz et al. (2003) investigated the relationship between the spiral arms of M81 and its X-ray source population, finding strong correlation between spiral arm position and X-ray source density. They note that brighter sources tend to be closer to spiral arms, attributable to the brightest and shortest-lived HMXBs being close downstream from their spiral arms. More recently, Kuntz et al. (2016) performed a deep *Chandra* survey of M51. This study also finds that X-ray sources are concentrated in spiral arms, though the distances to spiral arm midpoints are not presented. Both studies also found a non-trivial population of supernova remnants contributing to the total X-ray source population.

We also computed the distribution of distances to the nearest spiral arm across all the simulations, shown in Figure 4.8, in order to compare with previous works that measured the distances to OB associations and star-forming complexes for HMXBs (Bodaghee et al., 2012;

Coleiro & Chaty, 2013). In these works, clustering distances between HMXBs and SFCs/OB associations were inferred from the critical points of the cumulative distribution of the distances to the nearest SFC/OB association. As discussed in Section 4.1.1, distances to OB associations and SF complexes are distinct from distances to the spiral arms themselves, and as such we might not expect HMXBs to have the same clustering distance to the spiral arm. The distribution of HMXB distances to the nearest spiral arm that we measure does not show a strong preference for the clustering sizes measured for OB associations or SF complexes in previous works. Using the `quantiles` method of the `pandas` package, we measure the 0.1, 0.5, 0.9, and 0.99 quantiles of the HMXB distribution to be at 127, 570, 1296, and 2340 pc, respectively. For the LMXB distribution, the 0.1, 0.5, 0.9, and 0.99 quantiles of the LMXB distribution are at 129, 599, 1085, and 1783 pc, respectively.

Given the substantial width of these distributions, it is difficult to determine a characteristic separation from the spiral arms. The interarm separation of a few kpc as set by the symmetric arms model means that, by construction, it is difficult to have an XRB more than a few kpc away from a spiral arm in face-on projection. Further, we have chosen to model the galaxy using a symmetric model fitted to observables in the Milky Way, which is a simple albeit potentially unrealistic choice. The primary advantage of this model is that it permits us to easily define inter-arm and arm regions for analysis of the locations of XRBs. In reality, the number of arms and the symmetry (e.g., are the four arms symmetric with each other or are there major/minor axes?) of these arms in the Milky Way is difficult to characterize (see Vallée 2017 and references therein), and discussion exists about which tracers to use and how far to project the model based on nearby observables. Future attempts to characterize the relationship between the Galaxy's spiral arms and its XRB population would be improved by the use of a model that relaxes this symmetry constraint.

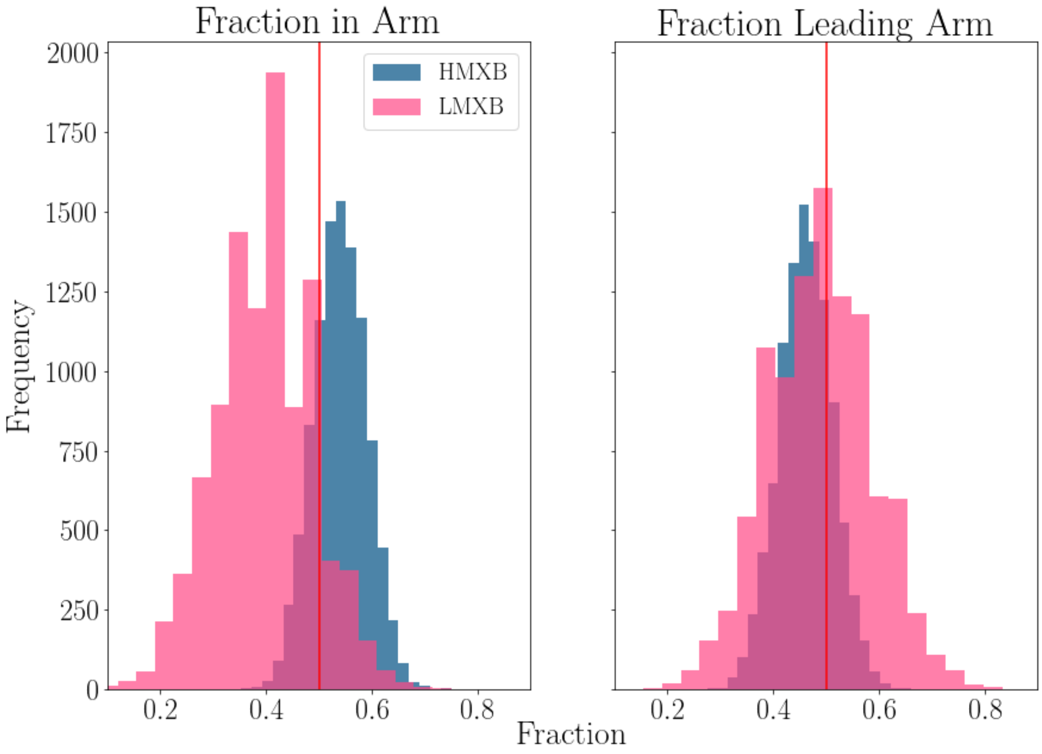

Figure 4.7: Distributions of population fraction correlated with spiral arms and leading edges for 10,000 realizations of the XRBs with Gaia candidate counterparts. The red line marks a fraction of 0.5, where the populations would be interpreted as being uncorrelated with the structure.

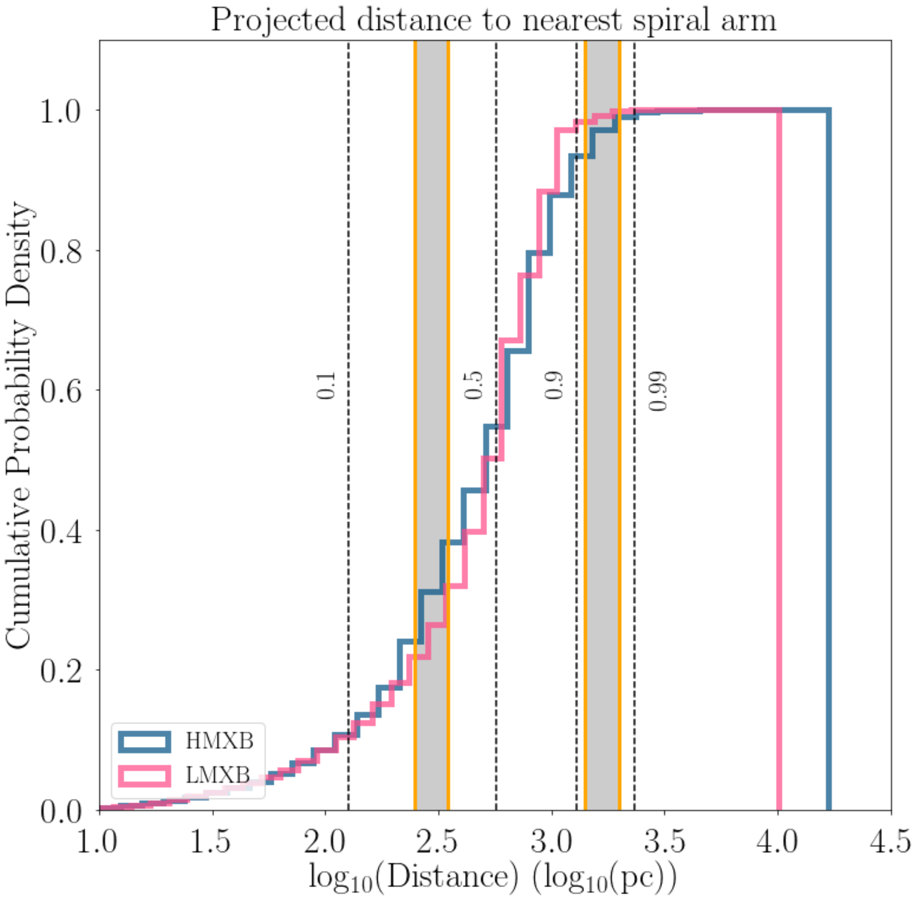

Figure 4.8: Cumulative distribution of XRB distances to the nearest spiral arm. We mark the characteristic clustering scales of HMXBs against OB associations and star-forming complexes measured by Bodaghee et al. 2012 and Coleiro & Chaty 2013 for reference using orange lines (for upper and lower limits) bounding grey regions. We also plot the 0.1, 0.5, 0.9, and 0.99 quantiles of the HMXB distribution for comparison.

Table 4.1: Properties of Gaia candidate counterparts to Galactic LMXBs

Names	RA	DEC	$P_{\mathrm{interloper}}$	Gaia ID	θ_{sep} ”	$m_{\mathrm{G,mean}}$ mag	d_{Gaia} kpc	d_{prev} kpc	d_{prev} Type	d_{prev} Ref
GRO J0422+32/V518 Per	04 21 42.790	+32 54 27.10	0.0100	172650748928103552	0.86	20.85 ± 0.05	···	···	···	···
Swift J061223.0+701243.9/-	06 12 22.600	+70 12 43.40	0.0001	1107229825742589696	0.29	21.00 ± 0.06	···	···	···	···
4U 0614+091/V1055 Ori	06 17 07.400	+09 08 13.60	0.0051	3328832132393159296	0.65	18.56 ± 0.02	$3.3^{+1.3}_{-2.4}$	3.0	burst	Brandt et al. (1992)
1A 0620-00/V616 Mon	06 22 44.503	-00 20 44.72	0.0070	3118721026600835328	0.73	17.52 ± 0.01	$1.6^{+0.4}_{-0.7}$	1.16	phot	Gelino et al. (2001)
4U 0919-54/*X	09 20 26.950	-55 12 24.70	0.0009	5310395631783100800	0.1	20.73 ± 0.02	···	5.4	burst	in’t Zand et al. (2005)
GS 1124-684/GU Mus	11 26 26.700	-68 40 32.60	0.0209	5234956524083372544	0.64	19.57 ± 0.01	$2.3^{+1.1}_{-3.1}$	5.5	phot	Orosz et al. (1996)
1A 1246-588/*X	12 49 39.364	-59 05 14.68	0.0002	6059778089610749440	0.04	20.495 ± 0.008	$2.0^{+1.2}_{-2.4}$	5.0	burst	Bassa et al. (2006)
4U 1456-32/V822 Cen	14 58 22.000	-31 40 08.00	0.0077	6205715168442046592	0.96	17.865 ± 0.005	$2.1^{+0.6}_{-1.2}$	1.3	burst	Kaluzienski et al. (1980)
3A 1516-569/BR Cir	15 20 40.900	-57 10 01.00	0.0763	5883218164517055488	0.87	17.92 ± 0.01	$6.2^{+2.0}_{-2.9}$	···	···	···
1E 1603.6+2600/UW CrB	16 05 45.820	+25 51 45.10	0.0018	1315375795016730880	0.75	19.67 ± 0.01	$2.1^{+0.7}_{-1.2}$	6.0	burst	Hakala et al. (2005)
H 1617-155/V818 Sco	16 19 55.070	-15 38 24.80	0.0001	4328198145165324800	0.22	12.48 ± 0.02	$2.13^{+0.21}_{-0.26}$	2.8	VLBAPLX	Bradshaw et al. (1999)
4U 1636-536/V801 Ara	16 40 55.500	-53 45 05.00	0.0819	5930753870442684544	0.87	18.27 ± 0.02	$4.4^{+1.6}_{-3.1}$	6.0	burst	Galloway et al. (2008a)
GRO J1655-40/V1033 Sco	16 54 00.137	-39 50 44.90	0.0022	5969790961312131456	0.15	16.224 ± 0.006	$3.3^{+0.7}_{-1.1}$	1.7	phot	Foellmi et al. (2006)
2A 1655+353/HZ Her	16 57 49.830	+35 20 32.60	< 0.0001	1338822021487330304	0.26	13.61 ± 0.02	$5.0^{+0.6}_{-0.7}$	6.6	SEDfit	Reynolds et al. (1997)
MXB 1659-298/V2134 Oph	17 02 06.500	-29 56 44.10	0.0517	6029391608332996224	0.48	19.44 ± 0.04	···	10.0	burst	Muno et al. (2001)
4U 1700+24/HD 154791	17 06 34.520	+23 58 18.60	< 0.0001	4571810378118789760	0.09	6.743 ± 0.001	$0.536^{+0.009}_{-0.009}$	0.42	phot	Masetti et al. (2002)
3A 1702-363/V1101 Sco	17 05 44.500	-36 25 23.00	0.0011	5976748056765619328	0.11	17.757 ± 0.008	$4.9^{+1.7}_{-2.9}$	···	···	···
SAX J1711.6-3808/-	17 11 37.100	-38 07 05.70	0.0299	5973177495780065664	0.99	21.05 ± 0.02	···	···	···	···
SAX J1711.6-3808/-	17 11 37.100	-38 07 05.70	0.0922	5973177495794993792	1.77	19.882 ± 0.007	$0.9^{+0.4}_{-2.9}$	···	···	···
4U 1724-307/Ter 2	17 27 33.300	-30 48 07.00	0.0347	4058208396397618688	0.34	18.20 ± 0.02	$6.6^{+2.8}_{-4}$	7.0	burst	Galloway et al. (2008a)
3A 1728-247/V2116 Oph	17 32 02.160	-24 44 44.00	0.0012	4110236324513030656	0.15	15.860 ± 0.009	$7.6^{+2.8}_{-4}$	4.5	phot	Chakrabarty & Roche (1997)
4U 1735-444/V926 Sco	17 38 58.300	-44 27 00.00	0.0742	5955379701104735104	0.96	17.77 ± 0.01	$5.6^{+2.1}_{-4}$	9.1	burst	Augusteijn et al. (1998)
SLX 1737-282/-	17 40 43.000	-28 18 11.90	0.0816	4060255373456473984	0.69	14.602 ± 0.003	$4.5^{+2.1}_{-4}$	7.5	burst	in’t Zand et al. (2002)
EXO 1747-214/star	17 50 24.520	-21 25 19.90	0.0067	4118590585673834624	0.17	20.24 ± 0.03	···	11.0	burst	Tomsick et al. (2005)
Swift J1753.5-0127/-	17 53 28.290	-01 27 06.22	< 0.0001	4178766135477201408	0.04	16.698 ± 0.009	$5.6^{+1.8}_{-2.8}$	6.0	A_V	Cadolle Bel et al. (2007)
4U 1755-33/V4134 Sgr	17 58 40.000	-33 48 27.00	0.0277	4042473487415175168	0.35	19.500 ± 0.007	···	6.5	phot	Wachter & Smale (1998)
2A 1822-371/V691 CrA	18 25 46.800	-37 06 19.00	0.0186	6728016172687965568	0.52	15.53 ± 0.02	$6.1^{+1.6}_{-2.7}$	2.5	SEDfit	Mason & Cordova (1982)
HETE J1900.1-2455/star	19 00 08.650	-24 55 13.70	0.0007	4074363039644919936	0.17	18.10 ± 0.01	$3.5^{+1.7}_{-3.5}$	5.0	burst	Kawai & Suzuki (2005)
4U 1908+005/V1333 Aql	19 11 16.000	+00 35 06.00	0.0638	4264296556603631872	0.87	18.901 ± 0.004	$3.0^{+1.3}_{-2.6}$	5.0	SEDfit	Rutledge et al. (2001)
4U 1916-05/V1405 Aql	19 18 47.870	-05 14 17.09	0.0066	4211396994895217152	0.37	20.92 ± 0.03	···	8.9	burst	Galloway et al. (2008a)
3A 1954+319/star	19 55 42.330	+32 05 49.10	0.0014	2034031438383765760	0.12	8.370 ± 0.002	$3.3^{+0.6}_{-1.0}$	1.7	phot	Masetti et al. (2006)
4U 2129+47/V1727 Cyg	21 31 26.200	+47 17 24.00	0.0106	1978241050130301312	0.53	17.600 ± 0.001	$1.75^{+0.26}_{-0.4}$	6.3	phot	Cowley & Schmidtke (1990)
GS 2023+338/V404 Cyg	20 24 03.830	+33 52 02.200	0.0058	2056188624872569088	0.24	17.19 ± 0.01	$2.1^{+0.4}_{-0.6}$	2.39	VLBIPLX	Miller-Jones et al. (2009)

First name in each column indicates the first name in the catalog, while the second name indicates the name of the optical counterpart (if any). Optical counterparts that have numbers/letters following a * refers to the corresponding object on the finding chart as described in the Liu catalogs. Optical counterparts with the name ”star” do not have a labelled object on their corresponding finding chart.

Table 4.2: Properties of Gaia candidate counterparts to Galactic HMXBs

Names	RA	DEC	$P_{\mathrm{interloper}}$	Gaia ID	θ_{sep} ”	$m_{\mathrm{G,mean}}$ mag	d_{Gaia} kpc	d_{prev} kpc	d_{prev} Type	d_{prev} Ref
1H 1555-552/HD 141926	15 54 21.80	-55 19 45.0	0.0610	5884544931471259136	0.74	8.680 ± 0.001	$1.35^{+0.08}_{-0.09}$	0.96	phot	Grillo et al. (1992)
IGR J16318-4848/*1	16 31 48.31	-48 49 00.7	0.0002	5940777877435137024	0.04	17.170 ± 0.002	$5.2^{+1.8}_{-2.7}$	0.9	SEDfit	Filliatre & Chaty (2004)
AX J1639.0-4642/-	16 39 05.40	-46 42 14.0	0.0149	5942638074996489088	0.4	19.800 ± 0.006	$4.0^{+1.7}_{-2.6}$	…	…	…
AX J1639.0-4642/-	16 39 05.40	-46 42 14.0	0.0358	5942638079315454976	0.64	20.61 ± 0.01	…	…	…	…
IGR J16465-4507/-	16 46 35.26	-45 07 04.5	0.0008	5943246345430928512	0.11	13.510 ± 0.001	$2.70^{+0.35}_{-0.5}$	12.5	phot	Smith (2004)
IGR J16479-4514/-	16 48 07.00	-45 12 05.8	0.0724	5940244030149933696	1.29	19.580 ± 0.005	$2.8^{+1.4}_{-2.5}$	…	…	…
4U 1700-37/HD 153919	17 03 56.80	-37 50 39.0	0.0154	5976382915813535232	0.33	6.400 ± 0.001	$1.75^{+0.19}_{-0.23}$	1.9	phot	Humphreys (1978)
XTE J1739-302/-	17 39 11.58	-30 20 37.6	0.0039	4056922105185686784	0.4	12.670 ± 0.001	$5.3^{+2.1}_{-4}$	2.3	phot	Negueruela et al. (2006)
IGR J17544-2619/*C1	17 54 25.28	-26 19 52.6	0.0006	4063908810076415872	0.1	11.670 ± 0.001	$2.66^{+0.33}_{-0.4}$	3.0	phot	Pellizza et al. (2006)
SAX J1819.3-2525/V4641 Sgr	18 19 21.48	-25 25 36.0	0.0168	4053096217067937664	0.28	18.82 ± 0.02	…	6.0	SEDfit	Chaty et al. (2003)
RX J1826.2-1450/LS 5039	18 26 15.06	-14 50 54.3	0.0004	4104196427943626624	0.08	10.8000 ± 0.0004	$1.96^{+0.19}_{-0.23}$	2.5	phot	Casares et al. (2005)
AX J1841.0-0536/-	18 41 00.43	-05 35 46.5	0.0017	4256500538116700160	0.09	12.940 ± 0.003	$7.6^{+2.2}_{-3.1}$	10.0	A_V	Bamba et al. (2001)
XTE J1901+014/star	19 01 39.90	+01 26 39.2	0.0045	4268294763113217152	0.6	19.450 ± 0.007	$2.2^{+1.1}_{-2.2}$	…	…	…
XTE J1906+09/star	19 04 47.48	+09 02 41.8	0.0014	4310649149314811776	0.23	19.73 ± 0.01	$2.8^{+1.4}_{-2.3}$	10.0	A_V	Marsden et al. (1998)
3A 1909+048/SS 433	19 11 49.60	+04 58 58.0	0.0102	4293406612283985024	0.56	12.63 ± 0.02	$3.8^{+0.8}_{-1.1}$	5.5	jetPM	Hjellming & Johnston (1981)
4U 1909+07/*A	19 10 48.20	+07 35 52.3	0.0028	4306419980916246656	0.62	20.170 ± 0.008	$2.6^{+1.3}_{-2.3}$	7.0	A_V	Wen et al. (2000)
IGR J19140+0951/-	19 14 04.20	+09 52 58.3	0.0056	4309253392325650176	0.41	18.200 ± 0.006	$2.8^{+1.3}_{-2.3}$	…	…	…
1H 1936+541/DM +53 2262	19 32 52.30	+53 52 45.0	0.0009	2136886799749672320	0.48	10.370 ± 0.002	$3.3^{+0.4}_{-0.5}$	…	…	…
XTE J1946+274/*A	19 45 39.30	+27 21 55.4	0.0422	2028089540103670144	0.76	15.7100 ± 0.0007	$12.6^{+2.9}_{-4}$	9.5	SEDfit	Wilson et al. (2003)
KS 1947+300/*3	19 49 30.50	+30 12 24.0	0.0359	2031938140034489344	0.57	20.48 ± 0.01	$3.1^{+2.0}_{-3.5}$	10.0	phot	Negueruela et al. (2003)
Cyg X-1/4U 1956+35/HD 226868	19 58 21.70	+35 12 06.0	0.0051	2059383668236814720	0.37	8.5200 ± 0.0008	$2.23^{+0.15}_{-0.18}$	2.14	cluster	Massey et al. (1995)
EXO 2030+375/*2	20 32 15.20	+37 38 15.0	0.0116	2063791369815322752	0.9	16.910 ± 0.003	$3.6^{+0.9}_{-1.3}$	7.1	A_V	Wilson et al. (2002)
RX J2030.5+4751/SAO 49725	20 30 30.80	+47 51 51.0	0.0125	2083644392294059520	0.54	9.0300 ± 0.0006	$2.49^{+0.16}_{-0.19}$	2.2	phot	Motch et al. (1997)
GRO J2058+42/star	20 58 47.50	+41 46 37.0	0.0048	2065653598916388352	0.45	14.190 ± 0.005	$8.0^{+0.9}_{-1.2}$	9.0	phot	Reig et al. (2005)
1H 2202+501/BD +49 3718	22 01 38.20	+50 10 05.0	0.0029	1979911002134040960	0.37	9.3000 ± 0.0004	$1.16^{+0.05}_{-0.05}$	0.7	HipPLX	Chevalier & Ilovaisky (1998)
4U 2206+543/BD +53 2790	22 07 56.20	+54 31 06.0	0.0086	2005653524280214400	0.52	9.7400 ± 0.0007	$3.34^{+0.32}_{-0.4}$	2.6	phot	Blay et al. (2006)
2S 0053+604/gamma Ca	00 56 42.50	+60 43 00.0	0.0041	426558460877467776	0.66	1.82 ± 0.01	…	0.19	HipPLX	Perryman et al. (1997)
2S 0114+650/V662 Cas	01 18 02.70	+65 17 30.0	0.0005	524924310153249920	0.17	10.520 ± 0.001	$6.6^{+1.1}_{-1.6}$	7.2	phot	Reig et al. (1996)
RX J0146.9+6121/LS I +61 235	01 47 00.20	+61 21 23.7	0.0001	511220031584305536	0.1	11.210 ± 0.002	$2.50^{+0.18}_{-0.21}$	2.3	phot	Coe et al. (1993)
IGR J01583+6713/-	01 58 18.44	+67 13 23.5	0.0009	518990967445248256	0.3	13.7000 ± 0.0007	$7.4^{+0.8}_{-1.1}$	…	…	…
1E 0236.6+6100/LS I +61 303	02 40 31.70	+61 13 46.0	0.0031	465645515129855872	0.48	10.390 ± 0.001	$2.45^{+0.21}_{-0.26}$	2.4	Kin	…
V 0332+53/BQ Cam	03 34 59.90	+53 10 24.0	0.0025	444752973131169664	0.71	14.220 ± 0.003	$5.1^{+0.8}_{-1.0}$	7.0	phot	Negueruela et al. (1999)
RX J0440.9+4431/LS V +44 17	04 40 59.30	+44 31 49.0	0.0014	252878401557369088	0.41	10.430 ± 0.001	$3.2^{+0.5}_{-0.6}$	3.2	phot	Motch et al. (1997)
EXO 051910+3737.7/V420 Aur	05 22 35.20	+37 40 34.0	0.0025	184497471323752064	0.52	7.220 ± 0.001	$1.29^{+0.09}_{-0.10}$	…	…	…
1A 0535+262/V725 Tau	05 38 54.60	+26 18 57.0	0.0011	3441207615229815040	0.37	8.680 ± 0.007	$2.13^{+0.21}_{-0.26}$	2.45	phot	Steele et al. (1998), Lyuty & Zaĭtseva (2000)
IGR J06074+2205/-	06 07 26.60	+22 05 48.3	0.0039	3423526544838563328	0.57	12.180 ± 0.001	$5.5^{+1.0}_{-1.5}$	…	…	…
SAX J0635.2+0533/-	06 35 18.29	+05 33 06.3	< 0.0001	3131755947406031104	0.15	12.510 ± 0.006	$5.7^{+1.3}_{-2.0}$	3.75	phot	Kaaret et al. (1999)
XTE J0658-073/[M81] I-33	06 58 17.30	-07 12 35.3	0.0004	3052677318793446016	0.2	12.030 ± 0.003	$5.1^{+0.9}_{-1.4}$	…	…	…
3A 0726-260/V441 Pup	07 28 53.60	-26 06 29.0	0.0019	5613494119544761088	0.3	11.620 ± 0.003	$9.5^{+2.1}_{-3.1}$	5.35	phot	Corbet & Mason (1984), Negueruela et al. (1996)
1H 0739-529/HD 63666	07 47 23.60	-53 19 57.0	0.0005	5489434710755238400	0.22	7.5200 ± 0.0004	$0.643^{+0.017}_{-0.018}$	0.52	HipPLX	Chevalier & Ilovaisky (1998)
4U 0900-40/HD 77581	09 02 06.90	-40 33 17.0	0.0044	5620657678322625920	0.46	6.720 ± 0.002	$2.42^{+0.16}_{-0.19}$	1.9	phot	Sadakane et al. (1985)
GRO J1008-57/star	10 09 46.90	-58 17 35.5	0.0052	5258414192353423360	0.52	13.900 ± 0.001	$3.6^{+0.4}_{-0.5}$	5.0	phot	Coe et al. (1994)

First name in each column indicates the first name in the catalog, while the second name indicates the name of the optical counterpart (if any). Optical counterparts that have numbers/letters following a * refers to the corresponding object on the finding chart as described in the Liu catalogs. Optical counterparts with the name ”star” do not have a labelled object on their corresponding finding chart.

Table 4.2: continued

Names	RA	DEC	$P_{\text{interloper}}$	Gaia ID	θ_{sep} ”	$m_{\text{G,mean}}$ mag	d_{Gaia} kpc	d_{prev} kpc	d_{prev} Type	d_{prev} Ref
1A 1118-615/Hen 3-640	11 20 57.20	-61 55 00.0	0.0055	5336957010898124160	0.25	11.6000 ± 0.0005	$2.93^{+0.22}_{-0.26}$	5.0	phot	Janot-Pacheco et al. (1981)
4U 1119-603/V779 Cen	11 21 15.10	-60 37 25.5	0.0021	5337498593446516480	0.14 0.14	12.890 ± 0.003	$6.4^{+1.0}_{-1.4}$	9.0	phot	Krzeminski (1974), Hutchings et al. (1979)
IGR J11215-5952/HD 306414	11 21 46.81	-59 51 47.9	0.0017	5339047221168787712	0.12	9.760 ± 0.001	$6.5^{+1.1}_{-1.5}$	8.0	phot	Negueruela et al. (2005)
IGR J11435-6109/-	11 44 10.70	-61 07 02.0	0.0363	5335022901224296064	0.64	13.000 ± 0.002	$3.9^{+1.1}_{-1.8}$	…	…	…
2S 1145-619/V801 Cen	11 48 00.00	-62 12 25.0	0.0023	5334823859608495104	0.18	8.630 ± 0.002	$2.23^{+0.16}_{-0.19}$	2.25	phot	Hutchings et al. (1981), Negueruela (1998)
1E 1145.1-6141/V830 Cen	11 47 28.60	-61 57 14.0	0.0594	5334851450481641088	0.64	12.280 ± 0.001	$9.1^{+1.6}_{-2.2}$	8.0	phot	Ilovaisky et al. (1982)
4U 1223-624/BP Cru ?	12 26 37.60	-62 46 13.0	0.0075	6054569565614460800	0.37	9.760 ± 0.001	$3.5^{+0.4}_{-0.5}$	5.0	phot	Kaper et al. (1995)
1H 1249-637/HD 110432	12 42 50.30	-63 03 31.0	0.0037	6055103928246312960	0.24	5.120 ± 0.002	$0.416^{+0.021}_{-0.023}$	0.3	HipPLX	Chevalier & Ilovaisky (1998)
1H 1253-761/HD 109857	12 39 14.60	-75 22 14.0	0.0003	5837600152935767680	0.16	6.5200 ± 0.0004	$0.2117^{+0.0014}_{-0.0014}$	0.24	HipPLX	Chevalier & Ilovaisky (1998)
4U 1258-61/V850 Cen	13 01 17.10	-61 36 07.0	0.0005	5863533199843070208	0.36	12.650 ± 0.003	$2.01^{+0.13}_{-0.15}$	2.4	phot	Parkes et al. (1980)
2RXP J130159.6-635806/-	13 01 58.70	-63 58 09.0	0.0029	5862285700835092352	0.22	17.340 ± 0.001	$5.5^{+1.7}_{-2.8}$	5.5	phot	Chernyakova et al. (2005)
2S 1417-624/*7	14 21 12.90	-62 41 54.0	0.0995	5854175187681966464	0.81	20.490 ± 0.006	$3.8^{+1.7}_{-2.7}$	6.2	phot	Grindlay et al. (1984)
4U 1538-52/QV Nor	15 42 23.30	-52 23 10.0	0.0267	5886085557746480000	0.72	13.190 ± 0.001	$6.6^{+1.4}_{-2.1}$	4.5	phot	Clark (2004), Reynolds et al. (1992)

4.4 Conclusions

- We have assembled the largest sample of Galactic X-ray binaries whose distances have all been measured using the same method, and hence have the same systematics and uniform presumed biases.
- Comparing XRB distances measured by Gaia to previous methods shows that measuring distances using Type I X-ray bursts appears to systematically overestimate distance. This suggests that assumptions about X-ray bursts, namely that bursting neutron stars consistently reach the Eddington luminosity, may need to be modified to use X-ray bursts as a distance estimator.
- We have also compared the positions of XRBs to the locations of the midpoints of spiral arms in the Milky Way. Galactic HMXBs in our sample show only a modest preference for being spatially co-located with spiral arms versus interarm regions, and show only a modest preference for being on the leading edge of spiral arms. This suggests that the delay time between star formation and HMXB formation/accretion beginning manifests itself observationally as a spatial separation between HMXBs and spiral arms due to the pattern speed of spiral arm rotation. Other possible explanations for this effect are scattering due to natal HMXB kicks or the possibility of star formation occurring closer to the midpoint of the arm than the leading edge.

Further releases of Gaia will hopefully yield additional Gaia candidate counterparts for Galactic XRBs, particularly for the intrinsically optically fainter LMXBs. For objects with identified Gaia candidate counterparts, smaller distance uncertainties will be possible using an improved baseline in DR3 and subsequent releases. A major limitation of our sample is the small sample size yielded from the Liu catalogs. The *Chandra* Source Catalog provides an excellent foundation for studying the Galactic X-ray sky in the *Chandra* era, but at present it has not been data-mined to make a Milky Way-specific catalogue as a potential successor to the Liu catalogs. Our knowledge of the Galactic XRB source population can be improved through future all-sky surveys, such as the planned eROSITA mission (Merloni et al., 2012). This mission, designed as a successor to the ROSAT mission, will survey the sky at approximately 20 times the sensitivity of ROSAT in soft X-rays (0.5 - 2.0 keV), while providing the first imaging survey of the sky in hard X-rays (2 - 10 keV). The on-axis angular resolution of this telescope is expected to be comparable to that of *XMM-Newton*. Aside from distance errors, a main source of uncertainty in our analysis is the low number of XRBs with Gaia counterparts. An improved all-sky survey will allow us to find Gaia counterparts to an X-ray catalogue that is more up-to-date and is has more uniform systematics, enhancing our understanding of how XRBs correlate to Galactic structure.

Bibliography

Augusteijn, T., van der Hooft, F., de Jong, J. A., van Kerkwijk, M. H., & van Paradijs, J. 1998, A&A, 332, 561

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, ArXiv e-prints. `https://arxiv.org/abs/1804.10121`

Bamba, A., Yokogawa, J., Ueno, M., Koyama, K., & Yamauchi, S. 2001, Publications of the Astronomical Society of Japan, 53, 1179, doi: `10.1093/pasj/53.6.1179`

Basinska, E. M., Lewin, W. H. G., Sztajno, M., Cominsky, L. R., & Marshall, F. J. 1984, ApJ, 281, 337, doi: `10.1086/162103`

Bassa, C. G., Jonker, P. G., in't Zand, J. J. M., & Verbunt, F. 2006, A&A, 446, L17, doi: `10.1051/0004-6361:200500229`

Berghoefer, T. W., Schmitt, J. H. M. M., & Cassinelli, J. P. 1996, Astronomy and Astrophysics Supplement Series, 118, 481

Bhattacharyya, S. 2010, Advances in Space Research, 45, 949, doi: `10.1016/j.asr.2010.01.010`

Blay, P., Negueruela, I., Reig, P., et al. 2006, A&A, 446, 1095, doi: `10.1051/0004-6361:20053951`

Bodaghee, A., Tomsick, J. A., Rodriguez, J., & James, J. B. 2012, ApJ, 744, 108, doi: `10.1088/0004-637X/744/2/108`

Boroson, B., Kim, D.-W., & Fabbiano, G. 2011, ApJ, 729, 12, doi: `10.1088/0004-637X/729/1/12`

Boubert, D., Guillochon, J., Hawkins, K., et al. 2018, MNRAS, 479, 2789, doi: `10.1093/mnras/sty1601`

Bozzo, E., Bahramian, A., Ferrigno, C., et al. 2018, A&A, 613, A22, doi: `10.1051/0004-6361/201832588`

Bradshaw, C. F., Fomalont, E. B., & Geldzahler, B. J. 1999, ApJ, 512, L121, doi: `10.1086/311889`

Brandt, S., Castro-Tirado, A. J., Lund, N., et al. 1992, A&A, 262, L15

Brown, A. G. A., Blaauw, A., Hoogerwerf, R., de Bruijne, J. H. J., & de Zeeuw, P. T. 1999, in NATO Advanced Science Institutes (ASI) Series C, Vol. 540, NATO Advanced Science Institutes (ASI) Series C, ed. C. J. Lada & N. D. Kylafis, 411

Cadolle Bel, M., Ribó, M., Rodriguez, J., et al. 2007, ApJ, 659, 549, doi: `10.1086/512004`

Casares, J., Jonker, P. G., & Israelian, G. 2017, X-Ray Binaries, ed. A. W. Alsabti & P. Murdin (Springer International Publishing AG), 1499

Casares, J., Ribó, M., Ribas, I., et al. 2005, MNRAS, 364, 899, doi: `10.1111/j.1365-2966.2005.09617.x`

Chaboyer, B. 1999, in Astrophysics and Space Science Library, Vol. 237, Post-Hipparcos Cosmic Candles, ed. A. Heck & F. Caputo, 111

Chakrabarty, D., & Roche, P. 1997, ApJ, 489, 254, doi: `10.1086/304779`

Chaty, S., Charles, P. A., Martí, J., et al. 2003, MNRAS, 343, 169, doi: `10.1046/j.1365-8711.2003.06651.x`

Chernyakova, M., Lutovinov, A., Rodríguez, J., & Revnivtsev, M. 2005, MNRAS, 364, 455, doi: `10.1111/j.1365-2966.2005.09548.x`

Chevalier, C., & Ilovaisky, S. A. 1998, A&A, 330, 201. `https://arxiv.org/abs/astro-ph/9710008`

Clark, G. W. 1975, ApJ, 199, L143, doi: `10.1086/181869`

—. 2004, ApJ, 610, 956, doi: `10.1086/421764`

Coe, M. J., Everall, C., Norton, A. J., et al. 1993, MNRAS, 261, 599, doi: `10.1093/mnras/261.3.599`

Coe, M. J., Roche, P., Everall, C., et al. 1994, MNRAS, 270, L57, doi: `10.1093/mnras/270.1.L57`

Coleiro, A., & Chaty, S. 2013, ApJ, 764, 185, doi: `10.1088/0004-637X/764/2/185`

Corbet, R. H. D., & Mason, K. O. 1984, A&A, 131, 385

Cowley, A. P., & Schmidtke, P. C. 1990, AJ, 99, 678, doi: `10.1086/115363`

Dabringhausen, J., Kroupa, P., Pflamm-Altenburg, J., & Mieske, S. 2012, ApJ, 747, 72, doi: `10.1088/0004-637X/747/1/72`

Degenaar, N., Wijnands, R., Cackett, E. M., et al. 2012, A&A, 545, A49, doi: `10.1051/0004-6361/201219470`

Dhawan, V., Mirabel, I. F., Ribó, M., & Rodrigues, I. 2007, ApJ, 668, 430, doi: `10.1086/520111`

Erkal, D., Boubert, D., Gualandris, A., Evans, N. W., & Antonini, F. 2018, ArXiv e-prints. `https://arxiv.org/abs/1804.10197`

Filliatre, P., & Chaty, S. 2004, ApJ, 616, 469, doi: `10.1086/424869`

Foellmi, C., Depagne, E., Dall, T. H., & Mirabel, I. F. 2006, A&A, 457, 249, doi: `10.1051/0004-6361:20054686`

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016a, A&A, 595, A1, doi: `10.1051/0004-6361/201629272`

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016b, A&A, 595, A2, doi: `10.1051/0004-6361/201629512`

—. 2018, A&A, 616, A1, doi: `10.1051/0004-6361/201833051`

Galloway, D. K., Muno, M. P., Hartman, J. M., Psaltis, D., & Chakrabarty, D. 2008a, The Astrophysical Journal Supplement Series, 179, 360, doi: `10.1086/592044`

Galloway, D. K., Özel, F., & Psaltis, D. 2008b, MNRAS, 387, 268, doi: `10.1111/j.1365-2966.2008.13219.x`

Galloway, D. K., Psaltis, D., Chakrabarty, D., & Muno, M. P. 2003, ApJ, 590, 999, doi: `10.1086/375049`

Gandhi, P., Rao, A., Johnson, M. A. C., Paice, J. A., & Maccarone, T. J. 2018, ArXiv e-prints. `https://arxiv.org/abs/1804.11349`

Gelino, D. M., Harrison, T. E., & Orosz, J. A. 2001, AJ, 122, 2668, doi: `10.1086/323714`

Gilfanov, M. 2004, MNRAS, 349, 146, doi: `10.1111/j.1365-2966.2004.07473.x`

González Hernández, J. I., Rebolo, R., Peñarrubia, J., Casares, J., & Israelian, G. 2005, A&A, 435, 1185, doi: `10.1051/0004-6361:20042453`

Grillo, F., Sciortino, S., Micela, G., Vaiana, G. S., & Harnden, F. R., J. 1992, The Astrophysical Journal Supplement Series, 81, 795, doi: `10.1086/191705`

Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2002, A&A, 391, 923, doi: `10.1051/0004-6361:20020826`

—. 2003, MNRAS, 339, 793, doi: `10.1046/j.1365-8711.2003.06224.x`

Grindlay, J. E., Petro, L. D., & McClintock, J. E. 1984, ApJ, 276, 621, doi: `10.1086/161650`

Hakala, P., Ramsay, G., Muhli, P., et al. 2005, MNRAS, 356, 1133, doi: `10.1111/j.1365-2966.2004.08543.x`

Harris, W. E. 1996, AJ, 112, 1487, doi: `10.1086/118116`

Hjellming, R. M., & Johnston, K. J. 1981, ApJ, 246, L141, doi: `10.1086/183571`

Hogg, D. W. 2018, ArXiv e-prints. `https://arxiv.org/abs/1804.07766`

Humphreys, R. M. 1978, The Astrophysical Journal Supplement Series, 38, 309, doi: `10.1086/190559`

Hutchings, J. B., Cowley, A. P., Crampton, D., van Paradijs, J., & White, N. E. 1979, ApJ, 229, 1079, doi: `10.1086/157042`

Hutchings, J. B., Crampton, D., & Cowley, A. P. 1981, AJ, 86, 871, doi: `10.1086/112959`

Ilovaisky, S. A., Chevalier, C., & Motch, C. 1982, A&A, 114, L7

in't Zand, J. J. M., Cumming, A., van der Sluys, M. V., Verbunt, F., & Pols, O. R. 2005, A&A, 441, 675, doi: `10.1051/0004-6361:20053002`

in't Zand, J. J. M., Verbunt, F., Kuulkers, E., et al. 2002, A&A, 389, L43, doi: `10.1051/0004-6361:20020631`

Janka, H.-T. 2013, MNRAS, 434, 1355, doi: `10.1093/mnras/stt1106`

Janot-Pacheco, E., Ilovaisky, S. A., & Chevalier, C. 1981, A&A, 99, 274

Jonker, P. G., Galloway, D. K., McClintock, J. E., et al. 2004, MNRAS, 354, 666, doi: `10.1111/j.1365-2966.2004.08246.x`

Kaaret, P., Piraino, S., Halpern, J., & Eracleous, M. 1999, ApJ, 523, 197, doi: `10.1086/307711`

Kaluzienski, L. J., Holt, S. S., & Swank, J. H. 1980, ApJ, 241, 779, doi: `10.1086/158388`

Kaper, L., Lamers, H. J. G. L. M., Ruymaekers, E., van den Heuvel, E. P. J., & Zuiderwijk, E. J. 1995, A&A, 300, 446. `https://arxiv.org/abs/astro-ph/9503003`

Kawai, N., & Suzuki, M. 2005, The Astronomer's Telegram, 534, 1

Koda, J., Scoville, N., Hasegawa, T., et al. 2012, ApJ, 761, 41, doi: `10.1088/0004-637X/761/1/41`

Krzeminski, W. 1974, ApJ, 192, L135, doi: `10.1086/181609`

Kuntz, K. D., Long, K. S., & Kilgard, R. E. 2016, ApJ, 827, 46, doi: `10.3847/0004-637X/827/1/46`

Kuulkers, E., den Hartog, P. R., in't Zand, J. J. M., et al. 2003, A&A, 399, 663, doi: `10.1051/0004-6361:20021781`

Lehmer, B. D., Alexander, D. M., Bauer, F. E., et al. 2010, ApJ, 724, 559, doi: `10.1088/0004-637X/724/1/559`

Lewin, W. H. G., van Paradijs, J., & Taam, R. E. 1993, Space Sci. Rev., 62, 223, doi: `10.1007/BF00196124`

Linares, M., Shahbaz, T., & Casares, J. 2018, ApJ, 859, 54, doi: `10.3847/1538-4357/aabde6`

Liu, Q. Z., van Paradijs, J., & van den Heuvel, E. P. J. 2006, A&A, 455, 1165, doi: `10.1051/0004-6361:20064987`

—. 2007, A&A, 469, 807, doi: `10.1051/0004-6361:20077303`

Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9, doi: `10.1051/0004-6361/201832964`

Lyuty, V. M., & Zaĭtseva, G. V. 2000, Astronomy Letters, 26, 9, doi: `10.1134/1.20364`

Marsden, D., Gruber, D. E., Heindl, W. A., Pelling, M. R., & Rothschild, R. E. 1998, ApJ, 502, L129, doi: `10.1086/311510`

Masetti, N., Orlandini, M., Palazzi, E., Amati, L., & Frontera, F. 2006, A&A, 453, 295, doi: `10.1051/0004-6361:20065025`

Masetti, N., Dal Fiume, D., Cusumano, G., et al. 2002, A&A, 382, 104, doi: `10.1051/0004-6361:20011543`

Mason, K. O., & Cordova, F. A. 1982, ApJ, 262, 253, doi: `10.1086/160416`

Massey, P., Johnson, K. E., & Degioia-Eastwood, K. 1995, ApJ, 454, 151, doi: `10.1086/176474`

Merloni, A., Predehl, P., Becker, W., et al. 2012, ArXiv e-prints. `https://arxiv.org/abs/1209.3114`

Miller-Jones, J. C. A., Jonker, P. G., Dhawan, V., et al. 2009, ApJ, 706, L230, doi: `10.1088/0004-637X/706/2/L230`

Mineo, S., Gilfanov, M., & Sunyaev, R. 2012, MNRAS, 419, 2095, doi: `10.1111/j.1365-2966.2011.19862.x`

Motch, C., Haberl, F., Dennerl, K., Pakull, M., & Janot-Pacheco, E. 1997, A&A, 323, 853. `https://arxiv.org/abs/astro-ph/9611122`

Muñoz-Darias, T., Casares, J., & Martínez-Pais, I. G. 2005, ApJ, 635, 502, doi: `10.1086/497420`

Muno, M. P., Chakrabarty, D., Galloway, D. K., & Savov, P. 2001, ApJ, 553, L157, doi: `10.1086/320682`

Muno, M. P., Pfahl, E., Baganoff, F. K., et al. 2005, ApJ, 622, L113, doi: `10.1086/429721`

Nättilä, J., Miller, M. C., Steiner, A. W., et al. 2017, A&A, 608, A31, doi: `10.1051/0004-6361/201731082`

Negueruela, I. 1998, A&A, 338, 505. `https://arxiv.org/abs/astro-ph/9807158`

Negueruela, I., Israel, G. L., Marco, A., Norton, A. J., & Speziali, R. 2003, A&A, 397, 739, doi: `10.1051/0004-6361:20021529`

Negueruela, I., Roche, P., Buckley, D. A. H., et al. 1996, A&A, 315, 160

Negueruela, I., Roche, P., Fabregat, J., & Coe, M. J. 1999, MNRAS, 307, 695, doi: `10.1046/j.1365-8711.1999.02682.x`

Negueruela, I., Smith, D. M., & Chaty, S. 2005, The Astronomer's Telegram, 470, 1

Negueruela, I., Smith, D. M., Harrison, T. E., & Torrejón, J. M. 2006, ApJ, 638, 982, doi: `10.1086/498935`

Orosz, J. A., Bailyn, C. D., McClintock, J. E., & Remillard, R. A. 1996, ApJ, 468, 380, doi: `10.1086/177698`

Parkes, G. E., Murdin, P. G., & Mason, K. O. 1980, MNRAS, 190, 537, doi: `10.1093/mnras/190.3.537`

Pellizza, L. J., Chaty, S., & Negueruela, I. 2006, A&A, 455, 653, doi: `10.1051/0004-6361:20054436`

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1997, A&A, 323, L49

Phillips, S. N., Shahbaz, T., & Podsiadlowski, P. 1999, MNRAS, 304, 839, doi: `10.1046/j.1365-8711.1999.02357.x`

Podsiadlowski, P., & Rappaport, S. 2000, ApJ, 529, 946, doi: `10.1086/308323`

Pooley, D., Lewin, W. H. G., Anderson, S. F., et al. 2003, ApJ, 591, L131, doi: `10.1086/377074`

Reig, P., Chakrabarty, D., Coe, M. J., et al. 1996, A&A, 311, 879

Reig, P., & Fabregat, J. 2015, A&A, 574, A33, doi: `10.1051/0004-6361/201425008`

Reig, P., Negueruela, I., Papamastorakis, G., Manousakis, A., & Kougentakis, T. 2005, A&A, 440, 637, doi: `10.1051/0004-6361:20052684`

Repetto, S., Davies, M. B., & Sigurdsson, S. 2012, MNRAS, 425, 2799, doi: `10.1111/j.1365-2966.2012.21549.x`

Reynolds, A. P., Bell, S. A., & Hilditch, R. W. 1992, MNRAS, 256, 631, doi: `10.1093/mnras/256.3.631`

Reynolds, A. P., Quaintrell, H., Still, M. D., et al. 1997, MNRAS, 288, 43, doi: `10.1093/mnras/288.1.43`

Rutledge, R. E., Bildsten, L., Brown, E. F., Pavlov, G. G., & Zavlin, V. E. 2001, ApJ, 559, 1054, doi: `10.1086/322361`

Sadakane, K., Hirata, R., Jugaku, J., et al. 1985, ApJ, 288, 284, doi: `10.1086/162791`

Smith, D. M. 2004, The Astronomer's Telegram, 338, 1

Steele, I. A., Negueruela, I., Coe, M. J., & Roche, P. 1998, MNRAS, 297, L5, doi: `10.1046/j.1365-8711.1998.01593.x`

Steiner, A. W., Heinke, C. O., Bogdanov, S., et al. 2018, MNRAS, 476, 421, doi: `10.1093/mnras/sty215`

Strohmayer, T., & Bildsten, L. 2006, in Compact stellar X-ray sources, ed. W. H. G. Lewin & M. van der Klis, Cambridge Astrophysics Series (Cambridge University Press), 113–156

Swartz, D. A., Ghosh, K. K., McCollough, M. L., et al. 2003, ApJS, 144, 213, doi: `10.1086/345084`

Tomsick, J. A., Gelino, D. M., & Kaaret, P. 2005, ApJ, 635, 1233, doi: `10.1086/497587`

Torrejón, J. M., & Orr, A. 2001, A&A, 377, 148, doi: `10.1051/0004-6361:20011070`

Tremmel, M., Fragos, T., Lehmer, B. D., et al. 2013, ApJ, 766, 19, doi: `10.1088/0004-637X/766/1/19`

Vallée, J. P. 2008, AJ, 135, 1301, doi: `10.1088/0004-6256/135/4/1301`

—. 2014, The Astronomical Journal, 148, 5

—. 2017, The Astronomical Review, 13, 113, doi: `10.1080/21672857.2017.1379459`

van Paradijs, J. 1978, Nature, 274, 650, doi: `10.1038/274650a0`

—. 1981, A&A, 101, 174

van Paradijs, J. 1998, in NATO Advanced Science Institutes (ASI) Series C, Vol. 515, NATO Advanced Science Institutes (ASI) Series C, ed. R. Buccheri, J. van Paradijs, & A. Alpar, 279

Verbunt, F. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 296, New Horizons in Globular Cluster Astronomy, ed. G. Piotto, G. Meylan, S. G. Djorgovski, & M. Riello, 245

Verbunt, F., & Hut, P. 1987, in IAU Symposium, Vol. 125, The Origin and Evolution of Neutron Stars, ed. D. J. Helfand & J.-H. Huang, 187

Verbunt, F., & Lewin, W. H. G. 2006, in Compact stellar X-ray sources, ed. W. H. G. Lewin & M. van der Klis, Cambridge Astrophysics Series (Cambridge University Press), 341–379

Verbunt, F., van Paradijs, J., & Elson, R. 1984, MNRAS, 210, 899, doi: `10.1093/mnras/210.4.899`

Wachter, S., & Smale, A. P. 1998, ApJ, 496, L21, doi: `10.1086/311242`

Wen, L., Remillard, R. A., & Bradt, H. V. 2000, ApJ, 532, 1119, doi: `10.1086/308604`

Wilson, C. A., Finger, M. H., Coe, M. J., Laycock, S., & Fabregat, J. 2002, ApJ, 570, 287, doi: `10.1086/339739`

Wilson, C. A., Finger, M. H., Coe, M. J., & Negueruela, I. 2003, ApJ, 584, 996, doi: `10.1086/345791`

Zhang, Z., Gilfanov, M., & Bogdán, Á. 2012, A&A, 546, A36, doi: `10.1051/0004-6361/201219015`

Chapter 5

Summary and Future Work

X-ray binaries are rare, unique objects which provide clues about fundamental physics, stellar evolution, and galactic evolution. Accretion, the main energy source for the radiation from X-ray binaries, is a fundamental process in astrophysics that describes the behaviour of a variety of systems at different astrophysical scales. The accretion in X-ray binaries is tremendously efficient at converting gravitational potential energy into high-energy radiation. In this thesis I have explored the nature of X-ray binaries by studying the X-ray binary population in three Local Group Galaxies.

X-ray binaries were discovered ~45 years ago during the beginning of X-ray astronomy with the launch of rocket-based experiments and eventually the *Uhuru* satellite (Giacconi et al., 1962, 1971). Many of the properties of X-ray binaries and their relationship to their host environments remain challenging to understand. X-ray observations are only possible from above the Earth's atmosphere, and the high energy of X-rays means that X-ray telescopes require careful design to ensure that X-rays are reflected rather than absorbed. Moreover, except in the case of the brightest X-ray sources, the number of photons measured from an X-ray source can be quite small – sometimes only a handful of source photons over multiple kiloseconds of observation. This means that X-ray telescopes yield noisier spectra than is typical of visible or near-visible light observations. X-ray telescopes also tend to have inherently lower angular resolution – with sub-arcsecond resolution only being achievable since the launch of the *Chandra* X-ray telescope (Weisskopf et al., 2000). Pinpointing an X-ray source and associating it with particular a object on the sky can therefore be difficult to achieve.

The best-studied populations of X-ray binaries are those in the Milky Way, other nearby galaxies, and their satellites. X-ray binaries can be studied in old populations with high and low stellar encounter rates in nearby globular clusters and dwarf galaxies, respectively. Nearby spiral galaxies similar to the Milky Way help us understand their evolution in the context of galactic structure and star formation rate. They also permit us to study the X-ray binary content of Milky Way-like galaxies at a more favourable viewing angle than the Milky Way itself, without many of the observational complications inherent in looking through the Milky Way's disk. So far, it is generally understood that the number of X-ray binaries with low-mass companions correlates with both the stellar mass and local stellar density of a population (Gilfanov, 2004; Bahramian et al., 2013). In addition, the star formation rate of a population is correlated with the number of X-ray binaries with high-mass companions (Grimm et al., 2003). The presence of X-ray sources whose X-ray flux is similar to those of X-ray binaries, including foreground stars, cataclysmic variables, background galaxies/AGNs, supernova remnants, and chromospherically active binaries means that the X-ray binary content of a population of X-ray sources can only be understood through careful study. Namely, understanding the X-ray binary content of a population requires the identification of features that can separate X-ray binaries from other X-ray objects, most often through identification of counterparts at longer wavelengths (Grimm et al., 2005). This may also be done through other empirical methods. X-ray hardness ratios, typically the ratio of flux in different bands across the X-ray range of a telescope, can be used as a parameter space wherein X-ray binaries are mostly separate from other objects (Prestwich et al., 2003). Additionally, the details of emission mechanisms of X-ray binaries means that caution must be taken when attempting to measure properties such as brightness, temperature, and distance.

5.1 Multiwavelength Survey of Sculptor Dwarf X-ray Sources

Nearby dwarf galaxies provide a way to investigate the X-ray binary content of an old stellar population that is expected to have a negligible number of binaries created through stellar encounters. Additionally, dwarf galaxies are sufficiently old that there should only be low-mass X-ray binaries, as any high-mass companions of high-mass X-ray binaries will have reached the endpoint of stellar evolution. Constraints on the low-mass X-ray binary production of these galaxies help us understand the production of X-ray binaries in both ordinary galaxies and globular clusters. In Chapter 2, we have performed a detailed study of the Sculptor Dwarf Spheroidal galaxy that follows up previously existing X-ray observations to determine if it has any bright X-ray binaries in the present day. We combined a new epoch of *Chandra* observations with *Spitzer* infrared imaging and *Gemini* optical imaging and spectroscopy. In addition to the nine previously studied sources, we identified four new bright X-ray sources in the new observations. Of these sources, seven are background active galactic nuclei, two are background galaxies, one is a bright foreground star, and three have uncertain identifications.

In the most conservative case, where we assume that sources that have uncertain identifications are unaffiliated with the galaxy, Sculptor Dwarf Spheroidal galaxy appears to lack any bright low-mass X-ray binaries in the present epoch. The absence of these bright low-mass X-ray binaries has two possible origins – either the dark matter halo of Sculptor is not sufficiently massive to retain low-mass X-ray binaries against the supernova kicks they receive when they form, or all of the low-mass X-ray binaries originally in Sculptor have burned out by the present day due to the compact object consuming its companion through accretion. Since many globular clusters also have old stellar populations, a consequence of the latter interpretation would be that all bright low-mass X-ray binaries observed in galactic globular clusters are created dynamically and are not primordial. Sculptor's lack of low-mass X-ray binaries also suggests that either it lacked globular clusters in the past (as it does in the present day), or if it did, they did not contribute any low-mass X-ray binaries to its current population. If Sculptor's X-ray binary population is representative of dwarf galaxies with similar masses and star formation histories, it is reasonable to conclude that dwarf galaxies do not appreciably contaminate their host galaxy with low-mass X-ray binaries when they strongly interact with them.

5.2 Random Forest Classification of M31 X-ray Sources

Identifying X-ray binaries in X-ray source populations in nearby galaxies is challenging in the absence of a longer-wavelength counterpart. Classification schemes that attempt to separate X-ray binaries from other kinds of X-ray sources (e.g., Prestwich et al., 2003) are not always reliable since other X-ray emitting objects have similar observed X-ray properties. In Chapter 3, we investigated whether commonly used supervised machine learning algorithms available in the Python package `sklearn` (Pedregosa et al., 2011) can provide a pathway for more accurately identifying new X-ray binary candidates. We considered a catalogue of 943 X-ray sources in the direction of the Andromeda Galaxy (M31), of which 163 had been previously classified as known X-ray binaries, active galactic nuclei (AGNs), foreground stars, or supernova remnants (Vulic et al., 2016). As input for the supervised machine learning algorithms, we used the broad-band *Chandra* 0.5 - 8.0 keV photon flux, the mean observed and mean incident energies, and fifteen photon flux ratios constructed by dividing the photon flux in different energy bands in the *Chandra* range by the broad-band photon flux. After testing a naive Bayes classifier, logistic regression, random forest classifier, multi-layer perceptron neural network, and linear support vector classifier, we find that the random forest algorithm has the best predictive power. Python and R implementations of the random forest algorithm gave similar performance.

We analyzed the dataset using both a multiclass approach and a binary approach, finding that if we consider only classifications on the basis of whether an object is an X-ray binary or not, we can achieve roughly 80 – 85% accuracy in identifying new X-ray binary candidates. The random forest classifier identifies 16 new X-ray binary candidates at 90% probability. We have crossmatched sources classified with sources in the Panchromatic Hubble Andromeda Treasury, finding five sources associated with PHAT-classified clusters, of which four are classified as X-ray binaries by the random forest algorithm. Finally, we find that X-ray bands that most strongly separate out X-ray binaries are typically narrower bands, such as the 1.7–2.8 keV and 0.5 - 1.0 keV photon fluxes. We conclude that supervised machine learning techniques can help identify strong X-ray binary candidates in cases where X-ray observations are all that is available.

5.3 Gaia distances to Galactic X-ray binaries

The relative proximity of the Milky Way's X-ray binaries means that they are the easiest to study in detail as we can see objects that are intrinsically much fainter than in other galaxies. Inside the Milky Way, however, relative distance uncertainties are much larger than for X-ray binaries in other populations. Modelling the emission from an X-ray binary in order to determine its intrinsic flux is more challenging than for ordinary stars, and X-ray binaries lack any known property that has allowed them to be used as standard candles. Additionally, X-ray binary emission can be variable and is dependent on a host of factors that are difficult to measure independently, such as the accretion state, mass transfer rate, or accretion geometry. As such, existing methods for measuring X-ray binary distance are widely-varying in their approach and dependent on modelling and uncertain systematics. For nearby objects, the parallax method can be used to geometrically determine distance based on the apparent shift of an object relative to the fixed background for observations taken at opposite locations in the Earth's orbit.

The purpose of the *Gaia* satellite is to catalogue parallaxes for approximately 100 million sources by the end of its five-year mission (Gaia Collaboration et al., 2016). In Chapter 4, we used Data Release 2 from the *Gaia* satellite (Gaia Collaboration et al., 2018) to look for *Gaia* counterparts to known X-ray binaries from the Liu et al. (2006) and Liu et al. (2007) catalogues of X-ray binaries. This provided us with the opportunity to create a sample of Galactic X-ray binaries whose distances are all determined by the same method. Of the ~300 sources in the two catalogues, we found 86 X-ray binaries which have a Gaia candidate counterpart, of which 32 are low-mass and 54 are high-mass X-ray binaries. We compared literature distances measured for these X-ray binaries to those derived using the parallaxes of the Gaia candidate counterpart and a Bayesian approach proposed by (Bailer-Jones et al., 2018), finding that *Gaia* candidate counterpart distances generally agree with distances from other parallax-based methods, but other methods tend to have significant scatter. Objects whose literature distance was determined using Type I X-ray bursts (Lewin et al., 1993) had systematically higher distances than the *Gaia* candidate counterpart. We conclude that this is tentative evidence that bursting neutron stars may not consistently reach the Eddington luminosity, as is typically assumed (Kuulkers et al., 2003).

We also compared the positions of X-ray binaries with *Gaia* candidate counterparts to the positions of the spiral arms in the Milky Way in a face-on projection (Vallée, 2008). The positions of Galactic low-mass X-ray binaries, were found to be modestly anti-correlated with spiral arms, though we could not reject the null hypothesis that they are uncorrelated with spiral arms. Measuring an anti-correlation is unexpected since low-mass X-ray binaries are a much older population relative to spiral arm rotation and they have received large supernova natal kicks that would displace them from the stellar population where they formed. The positions of Galactic high-mass X-ray binaries are modestly correlated with spiral arms and being located at the leading edge of their nearest spiral arm. However, our measurement of this correlation has sufficiently large uncertainties that we cannot reject the null hypothesis that high-mass X-ray binaries are uncorrelated with spiral arms. Other possible explanations are that the delay between star formation and high-mass X-ray binary formation is observationally manifested as a spatial separation between high-mass X-ray binaries and spiral arms. It is also possible that supernova kicks imparted to high-mass X-ray binaries, or star formation occurring closer to the midpoint of the spiral arm than the leading edge, may explain this weak correlation. We conclude that there is no evidence for a strong correlation between the positions of *Gaia* candidate counterparts to Galactic high-mass X-ray binaries and the spiral arms of the Galaxy.

5.4 Future Work

In this thesis, we have examined different aspects of the X-ray binary populations of three Local Group galaxies: the Sculptor Dwarf Spheroidal Galaxy, the Andromeda Galaxy, and the Milky Way itself. The bulk of this work consists of analysis and interpretation of observational data collected from a variety of sources. Future studies are therefore most straightforwardly done by obtaining additional observations and higher-quality observations than those used in this work.

This thesis provided a deep investigation into the bright X-ray source population of a dwarf galaxy. Our results show that this galaxy is essentially devoid of bright X-ray binaries, which could have implications for the observed bright X-ray binary population in other populations. Testing these implications in more detail would require studies of other nearby dwarf galaxies using a similar, detailed multiwavelength analysis. Such an investigation would ideally examine dwarf galaxies that do not have globular clusters, such as Draco and Ursa Major II, and those that do, such as the Fornax, Canis Major, and Sagittarius dwarf galaxies. In addition, our analysis focused on the observations of the bright X-ray sources in Sculptor, which are the most likely to be X-ray binaries. It is possible that there is an as-yet undetected population of quiescent X-ray binaries and cataclysmic variables. Deeper optical observations would be required to characterize the cataclysmic variable population or look for quiescent X-ray binaries.

We demonstrated, as proof of concept, that supervised machine learning techniques, specifically the random forest classifier, can help identify new X-ray binary candidates in the Andromeda Galaxy using sources which have been previously classified. This work could be enhanced by the addition of more classified examples of X-ray sources. For M31 specifically, we could expand the training set using more recent X-ray observations that overlap better with *HST* observations (Williams et al., 2018) and should yield additional classified examples. We would also like to improve our algorithm by adding classified examples from other nearby galaxies (e.g., Binder et al., 2015; Liu, 2011). This would make a larger, more generic training set that could be used to look for new X-ray binary candidates in a variety of populations, not just M31. A more detailed study could also modify the algorithm to account for measured uncertainties in flux ratios, extinction along different lines of sight, variability corrections, and the properties of optical counterparts.

Finally, we have investigated the relationship between X-ray binaries in the Milky Way and its spiral arm structure. Future *Gaia* data releases will hopefully yield additional counterparts to known Galactic X-ray binaries that have parallaxes. These data releases are expected to provide, through a longer baseline, more objects with a parallax, and lower uncertainties on previous parallax measurements. In addition, the X-ray binary catalogues used in our study are unfortunately out of date, despite being the most complete. Since these catalogues were last updated, a number of new Galactic X-ray binaries and X-ray binary candidates have been discovered. A future catalogue would take advantage of new X-ray telescopes that have been deployed in the interim, such as the Nuclear Spectroscopic Telescopic Array (*NuSTAR*) telescope, the *Astrosat* telescope, and the Neutron Star Interior Composition Explorer (*NICER*) (Gendreau et al., 2012; Harrison et al., 2013; Singh et al., 2014). Future X-ray missions, especially the planned *ATHENA* and *eROSITA*, will also compliment future studies that use *Gaia* distances to understand X-ray binaries properties. To move the work forward, one approach would be to data-mine the *Chandra* Source Catalog (Evans et al., 2010) to create an improved catalogue of Galactic X-ray binaries and X-ray binary candidates that can be cross-matched against current and future *Gaia* data releases.

Bibliography

Bahramian, A., Heinke, C. O., Sivakoff, G. R., & Gladstone, J. C. 2013, ApJ, 766, 136, doi: `10.1088/0004-637X/766/2/136`

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, ArXiv e-prints. `https://arxiv.org/abs/1804.10121`

Binder, B., Williams, B. F., Eracleous, M., et al. 2015, AJ, 150, 94, doi: `10.1088/0004-6256/150/3/94`

Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, ApJS, 189, 37, doi: `10.1088/0067-0049/189/1/37`

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: `10.1051/0004-6361/201629272`

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: `10.1051/0004-6361/201833051`

Gendreau, K. C., Arzoumanian, Z., & Okajima, T. 2012, in Proc. SPIE, Vol. 8443, Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray, 844313

Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. 1962, Physical Review Letters, 9, 439, doi: `10.1103/PhysRevLett.9.439`

Giacconi, R., Kellogg, E., Gorenstein, P., Gursky, H., & Tananbaum, H. 1971, ApJ, 165, L27, doi: `10.1086/180711`

Gilfanov, M. 2004, MNRAS, 349, 146, doi: `10.1111/j.1365-2966.2004.07473.x`

Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2003, MNRAS, 339, 793, doi: `10.1046/j.1365-8711.2003.06224.x`

Grimm, H.-J., McDowell, J., Zezas, A., Kim, D.-W., & Fabbiano, G. 2005, ApJS, 161, 271, doi: `10.1086/468185`

Harrison, F. A., Craig, W. W., Christensen, F. E., et al. 2013, ApJ, 770, 103, doi: `10.1088/0004-637X/770/2/103`

Kuulkers, E., den Hartog, P. R., in't Zand, J. J. M., et al. 2003, A&A, 399, 663, doi: `10.1051/0004-6361:20021781`

Lewin, W. H. G., van Paradijs, J., & Taam, R. E. 1993, Space Sci. Rev., 62, 223, doi: `10.1007/BF00196124`

Liu, J. 2011, ApJS, 192, 10, doi: `10.1088/0067-0049/192/1/10`

Liu, Q. Z., van Paradijs, J., & van den Heuvel, E. P. J. 2006, A&A, 455, 1165, doi: `10.1051/0004-6361:20064987`

—. 2007, A&A, 469, 807, doi: `10.1051/0004-6361:20077303`

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Prestwich, A. H., Irwin, J. A., Kilgard, R. E., et al. 2003, ApJ, 595, 719, doi: `10.1086/377366`

Singh, K. P., Tandon, S. N., Agrawal, P. C., et al. 2014, in Proc. SPIE, Vol. 9144, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray, 91441S

Vallée, J. P. 2008, AJ, 135, 1301, doi: `10.1088/0004-6256/135/4/1301`

Vulic, N., Gallagher, S. C., & Barmby, P. 2016, MNRAS, 461, 3443, doi: `10.1093/mnras/stw1523`

Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, in Proc. SPIE, Vol. 4012, X-Ray Optics, Instruments, and Missions III, ed. J. E. Truemper & B. Aschenbach, 2–16

Williams, B. F., Lazzarini, M., Plucinsky, P., et al. 2018, ArXiv e-prints. `https://arxiv.org/abs/1808.10487`

Curriculum Vitae

Name:	Robin Arnason
Post-Secondary Education and Degrees:	University of Western Ontario, London, ON 2014-2018 Ph.D., Astronomy
	University of Alberta, Edmonton, AB 2013-2014 MSc, Physics
	University of Alberta, Edmonton, AB 2008-2013 BSc, Physics Specialization
Honours and Awards:	*W. H. Wehlau Award* University of Western Ontario, London, Canada (2018)
	NSERC Alexander Graham Bell Canada Graduate Scholarship University of Western Ontario, London, Canada (2016-2018)
	Faculty of Science Doctoral Scholarship University of Western Ontario, London, Canada (2014-2016)
	Queen Elizabeth II Master's Scholarship University of Alberta, Edmonton, Canada (2012-2014)
	Jason Lang Scholarship University of Alberta, Edmonton, Canada (2008-2012)
	Valedictorian St. Thomas Aquinas High School, Spruce Grove, Canada (2008)

Related Work Experience:	Graduate Teaching Assistant University of Western Ontario (2014-2016, 2018)
	Image Processing Consultant Independent Contractor for Interface Fluidics, Inc. (2018)
	Data Science/Machine Learning Engineering Intern Atomic.vc (Rested) (2017)

Publications:

Arnason, R., Sivakoff, G., Heinke, C., Cohn, H., Lugger, P., "*A Low-mass Main-sequence Star and Accretion Disk in the Very Faint X-Ray Transient M15 X-3*", 2015, ApJ, 807, 1

Tetarenko, B., Bahramian, A., **Arnason, R.**, Miller-Jones, J., Repetto, S., Heinke, C., Maccarone, T., Chomiuk, L., Sivakoff, G., Strader, J., Kirsten, F., Vlemmings, W. "*The First Low-mass Black Hole X-Ray Binary Identified in Quiescence Outside of a Globular Cluster* ", 2016, MNRAS, 825, 1

Arnason, R., Barmby, P., Bahramian, A., Maccarone, T., Zepf., S. "*Multiwavelength survey of X-ray sources in Sculptor Dwarf Spheroidal Galaxy*", Submitted to MNRAS, May 2018

Work in Preparation

Arnason, R., Barmby, P., Vulic, N., "*Identifying New X-ray Binary Candidates in M31 using Random Forest Classification*", in prep.

Arnason, R., Bahramian, A., Papei, H., Barmby, P., Gorski, M., "*Probing Galactic X-ray Binaries and Galactic Structure with Gaia DR2*", in prep.

Conference Talks

Canadian Astronomical Society Conference (CASCA), Edmonton, AB (2017)

- Arnason, R., Barmby, P., Vulic, N., 2017, "Machine Learning Classification of X-ray Sources in the Andromeda Galaxy "

Compact Objects in Michigan 5, Lansing, MI (2017)

- Arnason, R., Barmby, P., Vulic, N., 2017, "Machine Learning Classification of X-ray Sources in the Andromeda Galaxy "

Canadian Astronomical Society Conference (CASCA), Quebec City, QC (2014)

- Arnason, R., Sivakoff, G., Heinke, C., Cohn, H., Lugger, P., 2014, “Challenging accretion models for Very Faint X-ray Transients with M15 X-3”

Conference Posters

Canadian Astronomical Society Conference (CASCA), Winnipeg, MB (2016)

- Arnason, R., Barmby, P., Vulic, N., 2016, “Machine Learning Classification of X-ray Sources in the Andromeda Galaxy”